

Dynamic Quadratic Cheap Talk and Signaling Games

Serkan Saritař, *Student Member, IEEE*, Serdar Yüksel, *Member, IEEE*,
and Sinan Gezici, *Senior Member, IEEE*

Abstract

Simultaneous (Nash) and sequential (Stackelberg) equilibria of two-player dynamic quadratic cheap talk and signaling game problems are investigated under a perfect Bayesian formulation. For the dynamic scalar and multi-dimensional cheap talk, it is shown that the Nash equilibrium cannot be fully revealing whereas the Stackelberg equilibrium is always fully revealing. In addition, the final state Nash equilibria have to be essentially quantized when the source is scalar, and non-revealing for the multi-dimensional case. In the dynamic signaling game where the transmission of a Gauss-Markov source over a memoryless Gaussian channel is considered, affine policies constitute an invariant subspace under best response maps for both scalar and multi-dimensional sources under Nash equilibria; however, the Stackelberg equilibrium policies are always linear for scalar sources but may be non-linear for multi-dimensional sources. Under the Stackelberg setup, the conditions under which the equilibrium is non-informative are derived for scalar sources, and a dynamic programming solution is presented when the encoders are restricted to be linear for multi-dimensional sources.

Index Terms

Stochastic networked control, information theory, game theory, signaling games, cheap talk.

I. INTRODUCTION

Signaling games and cheap talk are concerned with a class of Bayesian games where an informed player (encoder or sender) transmits information to another player (decoder or receiver). In these

S. Saritař and S. Gezici are with the Department of Electrical and Electronics Engineering, Bilkent University, 06800, Ankara, Turkey. Emails: {serkan.gezici}@ee.bilkent.edu.tr. S. Yüksel is with the Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada, K7L 3N6. Email: yuksel@mast.queensu.ca.

This research was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the Scientific and Technological Research Council (TÜBİTAK) of Turkey.

Part of this work was presented at the 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 2016. Part of this work will be presented at 2017 American Control Conference (ACC), Seattle, WA, 2017.

problems, the objective functions of the players are not aligned unlike the ones in the classical communication problems. The cheap talk problem was studied by Crawford and Sobel [1], who obtained the surprising result that under some technical conditions on the cost functions, the cheap talk problem only admits equilibria that involve quantized encoding policies. This is in contrast with the case where the goals are aligned in classical communication and information theory.

In this paper, we build on our earlier work [2] where we considered static quadratic cheap talk and signaling games, and we extend the analysis to the dynamic case.

A. Literature Review

As noted in [2], for team problems, although it is difficult to obtain optimal solutions under general information structures, it is apparent in such problems that more information provided to any of the decision makers does not negatively affect the utility of the players. There is also a well-defined partial order of information structures as studied by Blackwell [3] (see also its multi-agent extensions in [4]). However, for general zero-sum or non-zero-sum game problems, informational aspects are very challenging to address; more information can have negative effects on some or even all of the players in a system, see e.g. [5]. Further intricacies on informational aspects in competitive setups have been discussed in [6], [7] and [8].

The cheap talk and signaling game problems are applicable in networked control systems when a communication channel exists among competitive and non-cooperative decision makers. For example, in a smart grid application, there may be strategic sensors in the system [9] that wish to change the equilibrium for their own interests through reporting incorrect measurement values, and enforcing their outcome to enhance its prolonged use in the system. One further area of application is recommender systems [10]. For further applications, see [9], [11]. All of these applications lead to a new framework where the value of information and its utilization have a drastic impact on the system under consideration.

There have been extensive contributions to cheap talk and signaling games. The reader is referred to [2] for detailed discussion and references. On the dynamic side, much of the literature has focused on sequential (Stackelberg) equilibria as we note below. A notable exception is [12], where the dynamic (multi-stage) extension of the setup of Crawford and Sobel is analyzed for a source which is a fixed random variable distributed according to some density on $[0, 1]$ (see Remark 2.2 for a detailed discussion on this very relevant paper). [13] investigates a Gaussian cheap talk game under the Stackelberg assumption with quadratic cost functions for a class of single- and multi-terminal setups. It is shown that the best response of the transmitter is linear in the single-terminal setup whereas the policy space of

the transmitter is assumed to be affine in the multi-terminal setup. In [14], the non-alignment between the cost functions of the encoder and the decoder is a function of a Gaussian random variable (r.v.) and secret to the decoder; whereas, it is fixed and known to the decoder in [1], [2]. It is shown that the Stackelberg equilibrium strategies are affine in the quadratic Gaussian cheap talk setup. The dynamic Gaussian signaling game is studied as an extension of [14] in [15] where the linearity of Stackelberg equilibria is studied. [16] considers the information design problem between an encoder and a decoder with non-aligned utility functions under the Stackelberg equilibrium. [17] studies the central scheduling problem of allocating channels as a signaling game problem between the base station and mobile stations under the Stackelberg assumption. [18] investigates a dynamic linear quadratic Gaussian game with asymmetric information and simultaneous moves, and it is shown that under certain conditions, players' strategies that are linear in their private types.

In our earlier work [2], we considered both (simultaneous) Nash equilibria and (sequential) Stackelberg equilibria of the setup of Crawford and Sobel, and provided extensions to multi-dimensional and noisy setups. We showed that for scalar sources, the quantized nature of all equilibrium policies holds under Nash equilibria, whereas policies are fully informative under Stackelberg equilibria. Static signaling games were also considered, where Nash and Stackelberg equilibria were studied. In this paper, we build on [2], and extend the analysis to the dynamic case.

B. Preliminaries

A static *cheap talk* problem can be formulated as follows: An informed player (encoder) knows the value of the \mathbb{M} -valued random variable M and transmits the \mathbb{X} -valued random variable X to another player (decoder), who generates his \mathbb{M} -valued optimal decision U upon receiving X . The policies of the encoder and decoder are assumed to be deterministic; i.e., $x = \gamma^e(m)$ and $u = \gamma^d(x) = \gamma^d(\gamma^e(m))$. The encoder's goal is to minimize

$$J^e(\gamma^e, \gamma^d) = \mathbb{E}[c^e(m, u)] ,$$

whereas, the decoder's goal is to minimize

$$J^d(\gamma^e, \gamma^d) = \mathbb{E}[c^d(m, u)]$$

by finding optimal policies γ^e and γ^d , respectively. If the transmitted signal x is also an explicit part of the cost function c^e or c^d , then the communication between the players is not costless and the formulation turns into a *signaling game* problem. Such problems are studied under the tools and concepts provided by *game theory* since the goals are not aligned. In the simultaneous game-play; i.e.,

the encoder and decoder announce their policies at the same time, a pair of policies $(\gamma^{*,e}, \gamma^{*,d})$ is said to be a (simultaneous) **Nash equilibrium** [19] if

$$\begin{aligned} J^e(\gamma^{*,e}, \gamma^{*,d}) &\leq J^e(\gamma^e, \gamma^{*,d}) \quad \forall \gamma^e \in \Gamma^e \\ J^d(\gamma^{*,e}, \gamma^{*,d}) &\leq J^d(\gamma^{*,e}, \gamma^d) \quad \forall \gamma^d \in \Gamma^d \end{aligned} \quad (1)$$

where Γ^e and Γ^d are the sets of all deterministic functions from \mathbb{M} to \mathbb{X} and from \mathbb{X} to \mathbb{M} , respectively. Similarly, in the sequential game-play; i.e., first the encoder announces his policy, then the decoder (accordingly) announces his policy, a pair of policies $(\gamma^{*,e}, \gamma^{*,d})$ is said to be a **Stackelberg equilibrium** [19] if

$$J^e(\gamma^{*,e}, \gamma^{*,d}(\gamma^{*,e})) \leq J^e(\gamma^e, \gamma^{*,d}(\gamma^e)) \quad \forall \gamma^e \in \Gamma^e \quad (2)$$

where $\gamma^{*,d}(\gamma^e)$ satisfies

$$J^d(\gamma^e, \gamma^{*,d}(\gamma^e)) \leq J^d(\gamma^e, \gamma^d(\gamma^e)) \quad \forall \gamma^d \in \Gamma^d.$$

If an equilibrium is achieved when $\gamma^{*,e}$ is non-informative (e.g., the transmitted message and the source are independent) and $\gamma^{*,d}$ uses only the prior information (since the received message is useless), then we call such an equilibrium a *non-informative (babbling) equilibrium*. The following is a useful observation, which follows from [1]:

Proposition 1.1: A non-informative (babbling) equilibrium always exists for the cheap talk game.

Heretofore, only *static (one-stage) games* are considered. If a game is played over a number of time periods, the game is called a *dynamic game*. Let $m_{[0,N-1]} = \{m_0, m_1, \dots, m_{N-1}\}$ be a collection of random variables to be encoded sequentially (causally) to a decoder. In the k -th stage of an N -stage game, the encoder knows the values of $\mathcal{I}_k^e = \{m_{[0,k]}, x_{[0,k-1]}\}$ with $\mathcal{I}_0^e = \{m_0\}$, and transmits x_k to the decoder who generates his optimal decision by knowing the values of $\mathcal{I}_k^d = \{x_{[0,k]}\}$. Thus, under the policies considered, $x_k = \gamma_k^e(\mathcal{I}_k^e)$ and $u_k = \gamma_k^d(\mathcal{I}_k^d)$. The encoder's goal is to minimize

$$J^e(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d) = \mathbb{E} \left[\sum_{k=0}^{N-1} c_k^e(m_k, u_k) \right], \quad (3)$$

whereas the decoder's goal is to minimize

$$J^d(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d) = \mathbb{E} \left[\sum_{k=0}^{N-1} c_k^d(m_k, u_k) \right] \quad (4)$$

by finding the optimal policy sequences $\gamma_{[0,N-1]}^e = \{\gamma_0^e, \gamma_1^e, \dots, \gamma_{N-1}^e\}$ and $\gamma_{[0,N-1]}^d = \{\gamma_0^d, \gamma_1^d, \dots, \gamma_{N-1}^d\}$, respectively. Using the encoder cost in (3) and the decoder cost in (4), the Nash equilibrium and the Stackelberg equilibrium for dynamic games can be defined similarly as in (1) and (2), respectively.

Under both equilibria concepts, we consider the setups where the decision makers act optimally for each history path of the game (available to each decision maker) and the updates are Bayesian; thus the equilibria are to be interpreted under a *perfect Bayesian equilibria* concept. Since we assume such a (perfect Bayesian) framework, the equilibria lead to sub-game perfection and each decision maker performs optimal Bayesian decisions for every realized play path. For example, more general Nash equilibrium scenarios such as non-credible threats [20] or equilibria that are not strong time-consistent [21], [4, Definition 2.4.1] may not be considered.

In this paper, the quadratic cost functions are assumed; i.e., $c_k^e(m_k, u_k) = (m_k - u_k - b)^2$ and $c_k^d(m_k, u_k) = (m_k - u_k)^2$ where b is the bias term as in [1] and [2].

C. Contributions

The main contributions of this paper can be summarized as follows:

- We show that in the dynamic cheap talk game under Nash equilibria, the last stage equilibria are quantized for independent and identically distributed (i.i.d.) and Markov sources with arbitrary conditional probability measures, and fully revealing equilibria cannot exist in general (see Remark 2.1), whereas the equilibrium must be fully revealing in the dynamic scalar cheap talk game under Stackelberg equilibria.
- We show that the equilibria are fully revealing in the dynamic multi-dimensional cheap talk under Stackelberg equilibria whereas the equilibrium cannot be fully revealing under Nash equilibria.
- For the dynamic signaling game under Nash equilibria, it is shown that the encoder (decoder) must be affine for an affine decoder (encoder); namely, affine policies constitute an invariant subspace under best response maps.
- Dynamic Stackelberg signaling equilibria for scalar Gauss-Markov sources and scalar Gaussian channels are always linear, which is not necessarily the case for multi-dimensional setups.
- Further, the conditions for the existence of informative Stackelberg equilibria are provided for scalar sources through information theoretic arguments. Finally, a dynamic programming formulation is presented for Stackelberg equilibria when the encoders are restricted to be linear for multi-dimensional setups.

II. DYNAMIC CHEAP TALK

For the purpose of illustration, the system model of the 2-stage dynamic cheap talk is depicted in Fig. 1-(a).

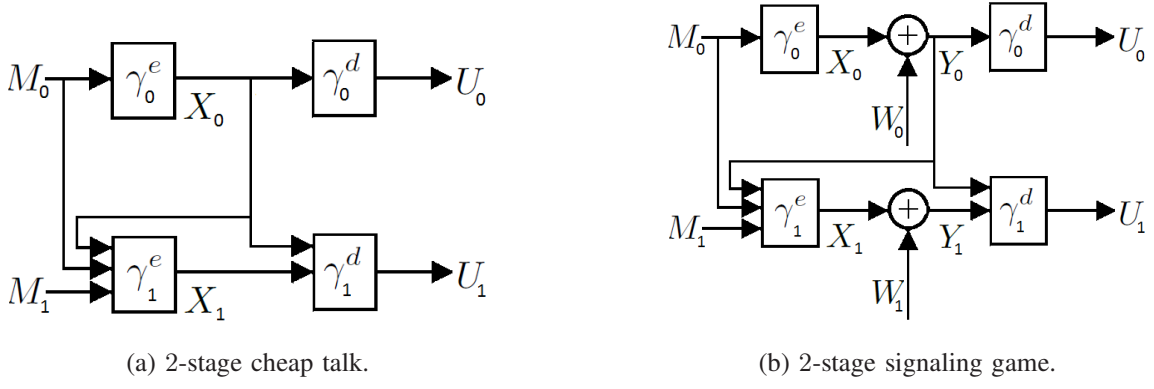


Fig. 1: General system models.

A. A Supporting Result : A Static Scalar Cheap Talk with Randomized Policies

To facilitate our analysis to handle certain intricacies that arise due to the dynamic setup in the paper, in the following we state that the result in [2, Theorem 3.2] also holds when the encoder is allowed to adapt randomized encoding policies by extending [1, Lemma 1] as follows:

Theorem 2.1: The conclusion of [2, Theorem 3.2], i.e., that an equilibrium policy is equivalent to a quantized policy, also holds if the policy space of the encoder is extended to the set of all stochastic kernels from \mathbb{M} to \mathbb{X} for any arbitrary source.¹ That is, even when the encoder is allowed to use private randomization, all equilibria are equivalent to those that are attained by quantized equilibria.

Proof: [1, Lemma 1] proves that all equilibria have finitely many partitions when the source has a bounded support. [2, Theorem 3.2] extends this result to a countable number of partitions for deterministic equilibria for any source with an arbitrary probability measure. The result follows by utilizing [2, Theorem 3.2] and [1, Lemma 1]. ■

Theorem 2.1 will be used crucially in the following analysis; since in a dynamic game, at a given time stage, the source variables from the earlier stages can serve as private randomness for the encoder.

B. Repeated i.i.d. Scalar Games: Nash Equilibria

In this part, the dynamic cheap talk game with an i.i.d. scalar source is analyzed.

Theorem 2.2:

¹Recall that P is a stochastic kernel from \mathbb{M} to \mathbb{X} if $P(\cdot|m)$ is a probability measure on $\mathcal{B}(\mathbb{X})$ for every $m \in \mathbb{M}$ and for every Borel $A \in \mathcal{B}(\mathbb{X})$, $P(A|\cdot)$ is a Borel measurable function of m .

- i) In the N -stage repeated cheap talk game, the equilibrium policies for the final stage must be quantized almost surely for any collection of policies $(\gamma_{[0,N-2]}^e, \gamma_{[0,N-2]}^d)$ and for any real-valued source model with arbitrary probability measure $P(dm_{N-1})$.
- ii) If the source m_k has a bounded support with a density, the first $N - 1$ stages cannot have fully revealing equilibria concurrently.

Proof: Here, we prove the results for the 2-stage setup, the extension to multiple stages is merely technical, as we comment on at the end of the proof.

- i) Let $c_1^e(m_1, u_1)$ be the second stage cost function of the encoder. Then the expected cost of the second stage encoder J_1^e can be written as follows:

$$\begin{aligned}
J_1^e &= \int P(dm_0, dm_1, dx_0, dx_1) c_1^e(m_1, u_1) \\
&= \int P(dx_0) P(dm_1|x_0) P(dm_0|m_1, x_0) P(dx_1|m_0, m_1, x_0) c_1^e(m_1, u_1) \\
&\stackrel{(a)}{=} \int P(dx_0) \int P(dm_1) P(dm_0|x_0) \mathbb{1}_{\{x_1=\gamma_1^e(m_0, m_1, x_0)\}} c_1^e(m_1, u_1) \\
&= \int P(dx_0) \int P(dm_1) P(dm_0|x_0) c_1^e(m_1, \gamma_1^d(x_0, \gamma_1^e(m_0, m_1, x_0)))
\end{aligned} \tag{5}$$

Here, (a) holds due to the i.i.d. source and the deterministic encoder assumptions. The inner integral of (5) can be considered as an expression for a given x_0 . Thus, given the second stage encoder and decoder policies $\gamma_1^e(m_0, m_1, x_0)$ and $\gamma_1^d(x_0, x_1)$, it is possible to define policies which are parametrized by the common information x_0 almost surely so that $\hat{\gamma}_{x_0}^e(m_0, m_1) = \gamma_1^e(m_0, m_1, x_0)$ and $\hat{\gamma}_{x_0}^d(x_1) = \gamma_1^d(x_0, x_1)$.

Now fix the first stage policies γ_0^e and γ_0^d . Suppose that the second stage encoder does not use m_0 ; i.e., $\hat{\gamma}_{x_0}^{e'}(m_1)$ is the policy of the second stage encoder. For the policies $\hat{\gamma}_{x_0}^{e'}(m_1)$ and $\hat{\gamma}_{x_0}^d(x_1)$, by using the second stage encoder cost function $F_{x_0}(m_1, u_1) \triangleq \mathbb{E}[(m_1 - u_1 - b)^2|x_0]$ and the bin arguments from [2, Theorem 3.2], it can be deduced that the equilibrium policies for the second stage must be quantized for any collection of policies (γ_0^e, γ_0^d) and for any given x_0 due to the continuity of $F_{x_0}(m_1, u_1)$ in m_1 . Now let the second stage encoder use m_0 ; i.e., $\hat{\gamma}_{x_0}^e(m_0, m_1)$ is the policy of the second stage encoder. Here, even if $\hat{\gamma}_{x_0}^e(m_0, m_1)$ is a deterministic policy, it can be regarded as an equivalent randomized encoder policy (as a stochastic kernel from \mathbb{M}_1 to \mathbb{X}_1) where m_0 is a real valued random variable independent of the source, m_1 . From Theorem 2.1, the equilibrium is achievable with an encoder policy which uses only m_1 ; i.e., $\hat{\gamma}_{x_0}^{e*}(m_1)$ is an encoder policy at the equilibrium and thus the equilibria are quantized.

For the N -stage game, the common information of the final stage encoder and decoder becomes

$x_{[0,N-2]}$ and $m_{[0,N-2]}$ is a vector valued random variable independent of the final stage source m_{N-1} .

- ii) Let two bins of the first stage equilibrium be \mathcal{B}_0^α and \mathcal{B}_0^β , and their encoding values be x_0^α and x_0^β , respectively. Also let m_0^α indicate any point in \mathcal{B}_0^α ; i.e., $m_0^\alpha \in \mathcal{B}_0^\alpha$. Similarly, let m_0^β represent any point in \mathcal{B}_0^β ; i.e., $m_0^\beta \in \mathcal{B}_0^\beta$. The decoder chooses an action $u_0^\alpha = \gamma_0^d(x_0^\alpha)$ when the encoder sends $x_0^\alpha = \gamma_0^e(m_0^\alpha)$ and an action $u_0^\beta = \gamma_0^d(x_0^\beta)$ when the encoder sends $x_0^\beta = \gamma_0^e(m_0^\beta)$ in order to minimize his total cost.

Let $F(m_0, x_0)$ be a cost function for the first stage encoder if it encodes message m_0 as x_0 . Since the second stage equilibrium cost does not depend on m_0 by the first part, $F(m_0, x_0)$ can be written as $F(m_0, x_0) = (m_0 - \gamma_0^d(x_0) - b)^2 + G(x_0)$ where $G(x_0) \triangleq \mathbb{E}_{m_1} \left[\left(m_1 - \gamma_1^{*,d}(x_0, \gamma_1^{*,e}(m_1, x_0)) - b \right)^2 \middle| x_0 \right]$ is the expected cost of the second stage encoder, and $\gamma_1^{*,e}$ and $\gamma_1^{*,d}$ are the second stage encoder and decoder policies at the equilibrium, respectively. Since under any equilibrium, the maximum number of bins is finite when the source has a bounded support, there are finitely many equilibria at the second stage which implies that the second stage encoder cost can take finitely many different values; i.e., $G(x_0)$ can take finitely many values.

Due to the equilibrium definitions from the view of the encoder, $F(m_0^\alpha, x_0^\alpha) < F(m_0^\alpha, x_0^\beta)$ and $F(m_0^\beta, x_0^\beta) < F(m_0^\beta, x_0^\alpha)$. These inequalities imply that

$$\begin{aligned} (m_0^\alpha - u_0^\alpha - b)^2 + G(x_0^\alpha) &< (m_0^\alpha - u_0^\beta - b)^2 + G(x_0^\beta) \\ (m_0^\beta - u_0^\beta - b)^2 + G(x_0^\beta) &< (m_0^\beta - u_0^\alpha - b)^2 + G(x_0^\alpha). \end{aligned} \tag{6}$$

In a fully revealing equilibrium, the encoder and the decoder policies are injective, thus these policies can be taken as identity functions; i.e., $x_0 = \gamma_0^e(m_0) = m_0$ and $u_0 = \gamma_0^d(x_0) = x_0 = m_0$. If we let $m_0^\alpha \rightarrow m_0^\beta$, then (6) becomes

$$\begin{aligned} G(m_0^\alpha) - G(m_0^\beta) &< (m_0^\alpha - m_0^\beta - b)^2 - b^2 \rightarrow 0 \\ G(m_0^\beta) - G(m_0^\alpha) &< (m_0^\beta - m_0^\alpha - b)^2 - b^2 \rightarrow 0. \end{aligned} \tag{7}$$

Thus if $m_0^\alpha \rightarrow m_0^\beta$ we must have $G(m_0^\alpha) \rightarrow G(m_0^\beta)$ which implies that $G(x_0)$ is continuous at $x_0 = m_0^\beta$. Since this is valid for any m_0^β and $G(x_0)$ can take finitely many values, $G(x_0)$ cannot have any jumps. Thus, it can be deduced that $G(x_0)$ is a constant function which is equivalent to say that if the first stage equilibrium is fully revealing, then the second stage equilibrium is constructed independently from the first stage equilibrium. Then (7) reduces to $b^2 < (m_0^\alpha - m_0^\beta - b)^2$ and $b^2 < (m_0^\beta - m_0^\alpha - b)^2$. After simplifications, it can be found that these inequalities are satisfied simultaneously if $|m_0^\alpha - m_0^\beta| > 2|b|$; however, this contradicts with the $m_0^\alpha \rightarrow m_0^\beta$ assumption.

Hence, the equilibrium cannot be fully informative at the first stage.

For an N -stage game, the analysis for the final stage works identically. ■

Remark 2.1: The boundedness assumption for the support of the measure $P(dm_k)$ can be relaxed for Theorem 2.2. In particular, a source with a probability measure $P(dm_k)$ that results in finitely many quantization bins in a static Nash equilibrium satisfies the conditions; e.g., when the random source has an exponential distribution ([22]).

C. Dynamic Game with a Scalar Markov Source: Nash Equilibria

In this part, the source M_k is assumed to be real valued Markovian for $k = 0, 1, \dots, N - 1$. The following result generalizes the first part of Theorem 2.2, which only considered i.i.d. sources.

Theorem 2.3: In the N -stage dynamic cheap talk game with a Markov source, the equilibrium policies for the final stage must be quantized almost surely for any collection of policies $(\gamma_{[0, N-2]}^e, \gamma_{[0, N-2]}^d)$ and for any real-valued source model with arbitrary probability measure.

Proof: Here, we prove the results for the 2-stage games as the extension is merely technical. The expected cost of the second stage encoder J_1^e can be written as follows similar to that in Theorem 2.2:

$$\begin{aligned} J_1^e &= \int P(dm_0, dm_1, dx_0, dx_1) c_1^e(m_1, u_1) \\ &= \int P(dx_0) \int P(dm_1|x_0) P(dm_0|m_1, x_0) c_1^e(m_1, \gamma_1^d(x_0, \gamma_1^e(m_0, m_1, x_0))) . \end{aligned}$$

After following similar arguments to those in the proof of Theorem 2.2, the second stage encoder policy becomes $\hat{\gamma}_{x_0}^e(m_0, m_1) \stackrel{(a)}{=} \hat{\gamma}_{x_0}^e(g(m_1, r), m_1) = \tilde{\gamma}_{x_0}^e(m_1, r)$ where (a) holds since any stochastic kernel from a complete, separable and metric space to another one, $P(dm_0|m_1)$, can be realized by some measurable function $m_0 = g(m_1, r)$ where r is a $[0, 1]$ -valued independent random variable (see Lemma 1.2 in [23], or Lemma 3.1 in [24]). Hence, the equilibria are quantized almost surely by Theorem 2.2. ■

Remark 2.2: A related setup has been studied in [12] where it has been shown that there can indeed be a fully revealing equilibrium if an individual source is transmitted repeatedly (thus the Markov source is a constant source). We note that there is no contradiction since for such a source, the terminal stage conditional measure can be made atomic via a careful construction of equilibrium policies for earlier time stages.

D. Dynamic Cheap Talk under Stackelberg Equilibria

In this part, the cheap talk game is analyzed under the Stackelberg assumption; i.e., the encoder knows the policy of the decoder. In this case, admittedly the problem is less interesting.

Theorem 2.4: For the 2-stage dynamic Stackelberg cheap talk game, any second stage encoder policy which uses (m_0, m_1, x_0) can be replaced, without any loss in performance, by one which uses only (m_1, x_0) .

Proof: For a fixed decoder policy $\gamma_1^d(x_0, x_1)$ where $x_1 = \gamma_1^e(m_0, m_1, x_0)$, the expected encoder cost can be written as

$$\begin{aligned} J_1^e &= \mathbb{E}[(m_1 - \gamma_1^d(x_0, \gamma_1^e(m_0, m_1, x_0)) - b)^2] \\ &= \int p(dm_0, dm_1) c(m_1, x_0, \gamma_1^e(m_0, m_1, x_0)) \\ &\geq \int p(dm_1) \tilde{c}(m_1, x_0, \gamma_1^{e,*}(m_1, x_0)). \end{aligned}$$

Here, the inequality follows from Blackwell's Irrelevant Information Theorem (see [25], [26]; or [4, p. 457]). Thus, the second stage encoder uses only m_1 and x_0 . ■

Remark 2.3: Theorem 2.4 is not valid if the simultaneous game-play is assumed. It is true that for a fixed decoder policy, any second stage encoder policy which uses (m_0, m_1, x_0) can be replaced, without any loss in performance, by one which uses only (m_1, x_0) . However, it is not possible to say anything about the equilibrium if the encoder uses (m_1, x_0) instead of (m_0, m_1, x_0) ; i.e., the fixed decoder policy may not be the optimal policy after the encoder changes his policy. To provide an interesting example: in the rock-paper-scissors game, the only equilibrium is a mixed equilibrium. For a random decoder policy, the encoder may switch to play always scissors instead of a mixed policy without changing his payoff. However, in this case, the optimal decoder policy is to choose the rock action always instead of a random policy.

Theorem 2.5: An equilibrium has to be fully revealing in the dynamic Stackelberg cheap talk game regardless of the source model.

Proof: We will use the properties of iterated expectations in the analysis. Recall that the total decoder cost is $J^d(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d) = \mathbb{E} \left[\sum_{k=0}^{N-1} (m_k - u_k)^2 \right]$. Considering the last stage, the goal of the decoder is to minimize $J_{N-1}^d(\gamma_{N-1}^e, \gamma_{N-1}^d) = \mathbb{E}[(m_{N-1} - u_{N-1})^2 | \mathcal{I}_{N-1}^d]$ by choosing the optimal action $u_{N-1}^* = \gamma_{N-1}^{*,d}(\mathcal{I}_{N-1}^d) = \mathbb{E}[m_{N-1} | \mathcal{I}_{N-1}^d]$. For the previous stage, the goal of the decoder is to minimize $J_{N-2}^d(\gamma_{N-1}^{*,e}, \gamma_{N-2}^e, \gamma_{N-1}^{*,d}, \gamma_{N-2}^d) = \mathbb{E}[(m_{N-2} - u_{N-2})^2 + J_{N-1}^{*,d}(\gamma_{N-1}^{*,e}, \gamma_{N-1}^{*,d}) | \mathcal{I}_{N-2}^d]$ by choosing the optimal action $u_{N-2}^* = \gamma_{N-2}^{*,d}(\mathcal{I}_{N-2}^d)$. Since $J_{N-1}^{*,d}(\gamma_{N-1}^{*,e}, \gamma_{N-1}^{*,d})$ is not affected by the choice of γ_{N-2}^d , the goal of the decoder is equivalent to the minimization of $\mathbb{E}[(m_{N-2} - u_{N-2})^2 | \mathcal{I}_{N-2}^d]$ at this stage.

Thus, the optimal policy is $u_{N-2}^* = \gamma_{N-2}^{*,d}(\mathcal{I}_{N-2}^d) = \mathbb{E}[m_{N-2}|\mathcal{I}_{N-2}^d]$. Similarly, since the actions taken by the decoder do not affect the future states and encoder policies, the optimal decoder actions can be found as $u_k^* = \gamma_k^{*,d}(\mathcal{I}_k^d) = \mathbb{E}[m_k|\mathcal{I}_k^d] = \mathbb{E}[m_k|x_{[0,k]}]$ for $k = 0, 1, \dots, N-1$.

Due to the Stackelberg assumption, the encoder knows that the decoder will use $u_k^* = \gamma_k^{*,d}(\mathcal{I}_k^d) = \mathbb{E}[m_k|\mathcal{I}_k^d]$ for each stage $k = 0, 1, \dots, N-1$. By using this assumption and the smoothing property of the expectation, the total encoder cost can be written as $J^e(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^{*,d}) = \mathbb{E} \left[\sum_{k=0}^{N-1} (m_k - u_k - b)^2 \right] = \mathbb{E} \left[\sum_{k=0}^{N-1} (m_k - u_k)^2 \right] + Nb^2$. Thus, as in the one-stage game setup [2, Theorem 3.3], the goals of the encoder and the decoder become essentially the same in the Stackelberg game setup, which effectively reduces the game setup to a team setup, resulting in fully informative equilibria; i.e., the encoder reveals all of his information. ■

E. Dynamic Multi-Dimensional Cheap Talk

In this section, Nash and Stackelberg equilibria of the dynamic multi-dimensional cheap talk are analyzed.

Since there may be discrete, non-discrete or even linear Nash equilibria in the static (one-stage) multi-dimensional cheap talk by [2, Theorem 3.4], the equilibrium policies are more difficult to characterize; however, we state the following:

Theorem 2.6: The Nash equilibrium cannot be fully revealing in the static (one-stage) multi-dimensional cheap talk when the source has positive measure for every non-empty open set.

Proof: Similar to the static scalar case [2, Theorem 3.2], in an equilibrium, define two cells C^α and C^β , any points in those cells as $\vec{m}^\alpha \in C^\alpha$ and $\vec{m}^\beta \in C^\beta$, and the actions of the decoder as \vec{u}^α and \vec{u}^β when the encoder transmits \vec{m}^α and \vec{m}^β , respectively. Let $F(\vec{m}, \vec{u}) \triangleq \|\vec{m} - \vec{u} - \vec{b}\|^2$. Due to the equilibrium definitions from the view of the encoder; $F(\vec{m}^\alpha, \vec{u}^\alpha) < F(\vec{m}^\alpha, \vec{u}^\beta)$ and $F(\vec{m}^\beta, \vec{u}^\beta) < F(\vec{m}^\beta, \vec{u}^\alpha)$. Hence, there exists a hyperplane defined by $F(\vec{z}, \vec{u}^\alpha) = F(\vec{z}, \vec{u}^\beta)$ which is equivalent to $\|(\vec{z} - \vec{b}) - \vec{u}^\alpha\|^2 = \|(\vec{z} - \vec{b}) - \vec{u}^\beta\|^2$. It can be seen that $\vec{z} - \vec{b}$ defines a hyperplane which is a perpendicular bisector of \vec{u}^α and \vec{u}^β ; i.e., the hyperplane defined by the points \vec{z} is a perpendicular bisector of \vec{u}^α and \vec{u}^β shifted by \vec{b} . The hyperplane defined by the points \vec{z} divides the space into two subspaces: let Z^α that contains \vec{u}^α and Z^β that contains \vec{u}^β be those subspaces. C^β and Z^α are disjoint subspaces since $F(\vec{z} + \Delta(\vec{u}^\beta - \vec{u}^\alpha), \vec{u}^\alpha) \geq F(\vec{z} + \Delta(\vec{u}^\beta - \vec{u}^\alpha), \vec{u}^\beta)$ for any positive scalar Δ . Similarly, C^α and Z^β are disjoint subspaces, too. Thus, the hyperplane defined by the points \vec{z} must lie between \vec{u}^α and \vec{u}^β which implies that the length of \vec{b} along the $\vec{d} \triangleq \vec{u}^\beta - \vec{u}^\alpha$ direction should not exceed half of the distance between \vec{u}^α and \vec{u}^β ; i.e., $\|\vec{b}_{\vec{d}}\| \leq \|\vec{d}\|/2$, where $\vec{b}_{\vec{d}}$ is the projection of \vec{b} along the direction

of \vec{d} . Since \vec{d} can be any vector at a fully revealing equilibrium by the assumption on the source, $\|\vec{b}_{\vec{d}}\| \leq \|\vec{d}\|/2$ cannot be satisfied unless $\vec{b} = \vec{0}$. Thus, there cannot be a fully revealing equilibrium in the static multi-dimensional cheap talk. ■

We can extend this result to the dynamic multi-dimensional cheap talk as follows:

Theorem 2.7: The final stage Nash equilibria cannot be fully revealing in the dynamic multi-dimensional cheap talk for i.i.d. and Markov sources when the conditional distribution $P(d\vec{m}_{N-1}|\vec{m}_{N-2})$ has positive measure for every non-empty open set.

Proof: The proof is the multi-dimensional extension of Theorem 2.2 for i.i.d. sources, and Theorem 2.3 for Markov sources. ■

Unlike the different characteristics between Nash equilibria of the dynamic scalar and multi-dimensional cheap talk, fully revealing characteristics of the Stackelberg equilibrium still hold for the dynamic multi-dimensional cheap talk, as for the scalar case:

Theorem 2.8: The Stackelberg equilibria in the dynamic multi-dimensional cheap talk can be obtained by extending its scalar case; i.e., it is unique and corresponds to a fully revealing encoder policy as in the scalar case.

Proof: Similar to the scalar case in Theorem 2.5, the optimal decoder actions are $\vec{u}_k^* = \mathbb{E}[\vec{m}_k|\vec{x}_{[0,k]}]$ for $k = 0, 1, \dots, N-1$. Then the total encoder cost becomes $J^e(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d) = \mathbb{E} \left[\sum_{k=0}^{N-1} \|\vec{m}_k - \vec{u}_k^*\|^2 \right] + N\|\vec{b}\|^2$, which effectively reduces the game setup to a team setup, resulting in fully informative equilibria. ■

As in the scalar case, the equilibria under the Nash and Stackelberg assumptions are drastically different: There cannot be fully revealing Nash equilibria in the dynamic multi-dimensional cheap talk whereas the equilibrium is always fully revealing under the Stackelberg assumption.

III. DYNAMIC QUADRATIC GAUSSIAN SIGNALING GAMES FOR SCALAR GAUSS-MARKOV SOURCES

The dynamic signaling game setup is similar to the dynamic cheap talk setup except that there exists an additive Gaussian noise channel between the encoder and decoder at each stage, and the encoder has a *soft* power constraint. For the purpose of illustration, the system model of the 2-stage dynamic signaling game is depicted in Fig. 1-(b).

Here, the source is assumed to be a Markov source with an initial Gaussian distribution; i.e., $M_0 \sim \mathcal{N}(0, \sigma_{M_0}^2)$ and $M_{k+1} = gM_k + V_k$ where $g \in \mathbb{R}$ and $V_k \sim \mathcal{N}(0, \sigma_{V_k}^2)$ is an i.i.d. Gaussian noise sequence for $k = 0, 1, \dots, N-2$. The channels between the encoder and the decoder are assumed to be i.i.d. additive Gaussian channels; i.e., $W_k \sim \mathcal{N}(0, \sigma_{W_k}^2)$, and W_k and V_l are independent for

$k = 0, 1, \dots, N - 1$ and $l = 0, 1, \dots, N - 2$. Since the messages transmitted by the encoder and received by the decoder are not the same due to the noisy channel, the information available to the encoder and the decoder slightly changes compared to that in the cheap talk setup. In the k -th stage of the N -stage game, the encoder knows the values of $\mathcal{I}_k^e = \{m_{[0,k]}, y_{[0,k-1]}\}$ (a noiseless feedback channel is assumed) and the decoder knows the values of $\mathcal{I}_k^d = \{y_{[0,k]}\}$ with $y_k = x_k + w_k$. Thus, under the policies considered, $x_k = \gamma_k^e(\mathcal{I}_k^e)$ and $u_k = \gamma_k^d(\mathcal{I}_k^d)$. The encoder's goal is to minimize

$$J^e(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d) = \mathbb{E} \left[\sum_{k=0}^{N-1} c_k^e(m_k, x_k, u_k) \right],$$

whereas, the decoder's goal is to minimize

$$J^d(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d) = \mathbb{E} \left[\sum_{k=0}^{N-1} c_k^d(m_k, u_k) \right].$$

by finding the optimal policy sequences $\gamma_{[0,N-1]}^e$ and $\gamma_{[0,N-1]}^d$, respectively. The cost functions are modified as $c_k^e(m_k, x_k, u_k) = (m_k - u_k - b)^2 + \lambda x_k^2$ and $c_k^d(m_k, u_k) = (m_k - u_k)^2$. Note that a power constraint with an associated multiplier is appended to the cost function of the encoder, which corresponds to power limitation for transmitters in practice. If $\lambda = 0$, this corresponds to the setup with no power constraint at the encoder.

A. Dynamic Nash Equilibria for Scalar Gauss-Markov Sources

In dynamic scalar signaling games, affine policies constitute an invariant subspace under best response maps for Nash equilibria which is stated as follows:

Theorem 3.1:

- i) If the encoder uses affine policies at all stages, then the decoder will also be affine at all stages.
- ii) If the decoder uses affine policies at all stages, then the encoder will also be affine at all stages.

Proof: A more general result is presented in Theorem 4.1. ■

Note that Theorem 3.1 does not lead to any conclusions about the informativeness of the equilibrium. Before delving into the informative equilibrium analysis of the 2-stage signaling game, the analysis of the static signaling game from [2] is refined as follows:

Theorem 3.2: There exists informative affine equilibria in the static signaling game if and only if $\frac{\sigma_M^2 - 2b^2 - \sqrt{\sigma_M^2} \sqrt{\sigma_M^2 - 4b^2}}{2\sigma_W^2} < \lambda < \frac{\sigma_M^2 - 2b^2 + \sqrt{\sigma_M^2} \sqrt{\sigma_M^2 - 4b^2}}{2\sigma_W^2}$ and $\sigma_M^2 \geq 4b^2$.

Proof: See Appendix A. ■

Next, the 2-stage signaling game is considered.

Theorem 3.3: For the 2-stage signaling game setup under affine encoder and decoder assumptions,

- 1) if $\lambda > \max \left\{ \frac{(g^2+1)\sigma_{M_0}^2}{\sigma_{W_0}^2}, \frac{\sigma_{M_1}^2}{\sigma_{W_1}^2} \right\}$, then there does not exist an informative affine equilibrium.
- 2) if $\min \left\{ \frac{(g^2+1)\sigma_{M_0}^2}{\sigma_{W_0}^2}, \frac{\sigma_{M_1}^2}{\sigma_{W_1}^2} \right\} < \lambda \leq \max \left\{ \frac{(g^2+1)\sigma_{M_0}^2}{\sigma_{W_0}^2}, \frac{\sigma_{M_1}^2}{\sigma_{W_1}^2} \right\}$, then;
 - a) for $\frac{(g^2+1)\sigma_{M_0}^2}{\sigma_{W_0}^2} < \frac{\sigma_{M_1}^2}{\sigma_{W_1}^2}$, the equilibrium is informative if and only if $\max \left\{ \frac{\sigma_{M_1}^2 - 2b^2 - \sqrt{\sigma_{M_1}^2(\sigma_{M_1}^2 - 4b^2)}}{2\sigma_{W_1}^2}, (g^2+1) \frac{\sigma_{M_0}^2}{\sigma_{W_0}^2} \right\} < \lambda < \frac{\sigma_{M_1}^2 - 2b^2 + \sqrt{\sigma_{M_1}^2(\sigma_{M_1}^2 - 4b^2)}}{2\sigma_{W_1}^2}$ and $\sigma_{M_1}^2 \geq 4b^2$.
 - b) for $\frac{\sigma_{M_1}^2}{\sigma_{W_1}^2} < \frac{(g^2+1)\sigma_{M_0}^2}{\sigma_{W_0}^2}$, the second stage message m_1 is not used in the game.

Proof: See Appendix B. ■

The analysis in Theorem 3.3 can be carried over to the N -stage signaling game; however, for an N -stage problem, this would involve $(3N^2 + 5N)/2$ equations and as many unknowns.

B. Dynamic Stackelberg Equilibria for Scalar Gauss-Markov Sources

In this section, the signaling game is analyzed under the Stackelberg concept. The equilibrium drastically changes under the Stackelberg assumption as shown below:

Theorem 3.4: An equilibrium has to be always linear in the dynamic Stackelberg signaling game. Further, there does not exist an informative (affine or non-linear) equilibrium in the N -stage dynamic scalar signaling game under the Stackelberg assumption; i.e., the only equilibrium is the non-informative one, if $\lambda \geq \max_{k=0,1,\dots,N-1} \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \sum_{i=0}^{N-k-1} g^{2i}$.

Proof: Similar to the dynamic Stackelberg cheap talk analysis in Theorem 2.5, the optimal decoder actions can be found as $u_k^* = \gamma_k^{*,d}(\mathcal{I}_k^d) = \mathbb{E}[m_k | \mathcal{I}_k^d] = \mathbb{E}[m_k | y_{[0,k]}]$ for $k = 0, 1, \dots, N-1$.

Due to the Stackelberg assumption, the encoder knows that the decoder will use $u_k^* = \gamma_k^{*,d}(\mathcal{I}_k^d) = \mathbb{E}[m_k | \mathcal{I}_k^d]$ for each stage $k = 0, 1, \dots, N-1$. Based on this assumption and the smoothing property of the expectation, the total encoder cost can be written as $J^e(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d) = \mathbb{E} \left[\sum_{k=0}^{N-1} (m_k - u_k - b)^2 + \lambda x_k^2 \right] = \mathbb{E} \left[\sum_{k=0}^{N-1} \mathbb{E}[(m_k - \mathbb{E}[m_k | \mathcal{I}_k^d])^2 + b^2 + \lambda x_k^2 | \mathcal{I}_k^d] \right]$. This problem is an instance of problems studied in [27] and [4, Theorem 11.3.1], and can be reduced to a team problem where both the encoder and the decoder are minimizing the same cost. The linearity of the optimal encoder and decoder can be deduced from [27].

For the second part of the proof, the lower bound for the encoder cost will be obtained and analyzed. From the chain rule, $I(m_k; y_{[0,k]}) = I(m_k; y_{[0,k-1]}) + I(m_k; y_k | y_{[0,k-1]})$. By following similar arguments to those in [28] and [4, Theorem 11.3.1],

$$\begin{aligned} I(m_k; y_k | y_{[0,k-1]}) &= h(y_k | y_{[0,k-1]}) - h(y_k | m_k, y_{[0,k-1]}) \\ &= h(y_k | y_{[0,k-1]}) - h(y_k | m_k, y_{[0,k-1]}, \gamma_k^e(m_k, y_{[0,k-1]})) \end{aligned}$$

$$\begin{aligned}
&= h(y_k|y_{[0,k-1]}) - h(y_k|\gamma_k^e(m_k, y_{[0,k-1]})) \leq h(y_k) - h(y_k|\gamma_k^e(m_k, y_{[0,k-1]})) \\
&= I(\gamma_k^e(m_k, y_{[0,k-1]}); y_k) = I(x_k; y_k) \leq \sup I(x_k; y_k) \\
&= \frac{1}{2} \log_2 \left(1 + \frac{P_k}{\sigma_{W_k}^2} \right) \triangleq \hat{C}_k \text{ where } P_k = \mathbb{E}[x_k^2].
\end{aligned}$$

It can be seen that $m_k - \mathbb{E}[m_k|m_{k-1}]$ is orthogonal to the random variables $m_{k-1}, y_{[0,k-1]}$ where $y_{[0,k-1]}$ is included due to the Markov chain $m_k \leftrightarrow m_{k-1} \leftrightarrow (y_{[0,k-1]})$. By using this orthogonality, it follows that

$$\begin{aligned}
\mathbb{E}[(m_k - \mathbb{E}[m_k|y_{[0,k-1]}])^2] &= \mathbb{E}[(m_k - \mathbb{E}[m_k|m_{k-1}])^2] + \mathbb{E}[(\mathbb{E}[m_k|m_{k-1}] - \mathbb{E}[m_k|y_{[0,k-1]}])^2] \\
&\stackrel{(a)}{=} \mathbb{E}[(m_k - \mathbb{E}[m_k|m_{k-1}])^2] + \mathbb{E}\left[(\mathbb{E}[m_k|m_{k-1}] - \mathbb{E}[\mathbb{E}[m_k|m_{k-1}, y_{[0,k-1]}]|y_{[0,k-1]}])^2\right] \\
&\stackrel{(b)}{=} \mathbb{E}[(m_k - \mathbb{E}[m_k|m_{k-1}])^2] + \mathbb{E}\left[(\mathbb{E}[m_k|m_{k-1}] - \mathbb{E}[\mathbb{E}[m_k|m_{k-1}]|y_{[0,k-1]}])^2\right] \\
&\stackrel{(c)}{=} \sigma_{V_{k-1}}^2 + g^2 \mathbb{E}[(m_{k-1} - \mathbb{E}[m_{k-1}|y_{[0,k-1]}])^2] \\
&\stackrel{(d)}{\geq} \sigma_{V_{k-1}}^2 + g^2 \sigma_{M_{k-1}}^2 2^{-2C_{k-1}}
\end{aligned} \tag{8}$$

where $C_k \triangleq \sup I(m_k; y_{[0,k]})$. Here, (a) holds due to the iterated expectation rule, (b) holds due to the Markov chain property, (c) holds since $\mathbb{E}[m_k|m_{k-1}] = \mathbb{E}[gm_{k-1} + v_{k-1}|m_{k-1}] = gm_{k-1}$, and (d) holds due to [4, Lemma 11.3.1]. From [4, Lemma 11.3.2], $I(m_k; y_{[0,k-1]})$ is maximized with linear policies, and the lower bound of (8), $\mathbb{E}[(m_k - \mathbb{E}[m_k|y_{[0,k-1]}])^2] \geq \sigma_{V_{k-1}}^2 + g^2 \sigma_{M_{k-1}}^2 2^{-2C_{k-1}} \triangleq \sigma_{M_k}^2 2^{-2\tilde{C}_k}$, is achievable through linear policies where $\sup I(m_k; y_{[0,k-1]}) \triangleq \tilde{C}_k = \frac{1}{2} \log_2 \left(\frac{\sigma_{M_k}^2}{\sigma_{V_{k-1}}^2 + g^2 \sigma_{M_{k-1}}^2 2^{-2C_{k-1}}} \right)$. Thus, we have the following recursion on upper bounds on mutual information for the N -stage dynamic signaling game:

$$\begin{aligned}
C_k &= \sup I(m_k; y_{[0,k]}) \\
&= \sup I(m_k; y_{[0,k-1]}) + \sup I(m_k; y_k|y_{[0,k-1]}) \\
&= \tilde{C}_k + \hat{C}_k \\
&= \frac{1}{2} \log_2 \left(\frac{\sigma_{M_k}^2}{\sigma_{V_{k-1}}^2 + g^2 \sigma_{M_{k-1}}^2 2^{-2C_{k-1}}} \right) + \frac{1}{2} \log_2 \left(1 + \frac{P_k}{\sigma_{W_k}^2} \right)
\end{aligned}$$

for $k = 1, 2, \dots, N-1$ with $C_0 = \frac{1}{2} \log_2 \left(1 + \frac{P_0}{\sigma_{W_0}^2} \right)$. Let the lower bound on $\mathbb{E}[(m_k - \mathbb{E}[m_k|y_{[0,k]}])^2]$ be Δ_k ; i.e., $\mathbb{E}[(m_k - \mathbb{E}[m_k|y_{[0,k]}])^2] \geq \sigma_{M_k}^2 2^{-2C_k} \triangleq \Delta_k$. Then the following recursion can be obtained for the N -stage dynamic signaling game:

$$\Delta_k = \frac{\sigma_{V_{k-1}}^2 + g^2 \Delta_{k-1}}{1 + \frac{P_k}{\sigma_{W_k}^2}} \text{ for } k = 1, 2, \dots, N-1$$

with $\Delta_0 = \frac{\sigma_{M_0}^2}{1 + \frac{P_0}{\sigma_{W_0}^2}}$. Since $\Delta_k = \sigma_{M_k}^2 2^{-2C_k}$ by definition, $\Delta_k \leq \sigma_{M_k}^2$ for $k = 0, 1, \dots, N-1$. In an equilibrium, since the decoder always chooses $u_k = \mathbb{E}[m_k|y_{[0,k]}]$ for $k = 0, 1, \dots, N-1$, the total encoder cost for the first stage can be lower bounded by $J_0^{e,lower} = \sum_{i=0}^{N-1} (\Delta_i + \lambda P_i + b^2)$. Now observe the following:

$$\frac{\partial \Delta_l}{\partial P_k} = \begin{cases} 0 & \text{if } l < k \\ g^2 \left(1 + \frac{P_l}{\sigma_{W_l}^2}\right)^{-1} \frac{\partial \Delta_{l-1}}{\partial P_k} - \frac{1}{\sigma_{W_l}^2} \frac{\partial P_l}{\partial P_k} \left(\sigma_{V_{l-1}}^2 + g^2 \Delta_{l-1}\right) \left(1 + \frac{P_l}{\sigma_{W_l}^2}\right)^{-2} & \text{if } l \geq k \end{cases}$$

where $\frac{\partial P_l}{\partial P_k} = 0$ for $l < k$ due to the information structure of the encoder. Then we obtain the following:

$$\begin{aligned} \frac{\partial J_0^{e,lower}}{\partial P_{N-1}} &= \lambda - \left(\sigma_{V_{N-2}}^2 + g^2 \Delta_{N-2}\right) \left(1 + \frac{P_{N-1}}{\sigma_{W_{N-1}}^2}\right)^{-2} \frac{1}{\sigma_{W_{N-1}}^2} \\ &\geq \lambda - \left(\sigma_{V_{N-2}}^2 + g^2 \sigma_{M_{N-2}}^2\right) \frac{1}{\sigma_{W_{N-1}}^2} \\ &= \lambda - \frac{\sigma_{M_{N-1}}^2}{\sigma_{W_{N-1}}^2}. \end{aligned}$$

If $\lambda > \frac{\sigma_{M_{N-1}}^2}{\sigma_{W_{N-1}}^2}$, then $\frac{\partial J_0^{e,lower}}{\partial P_{N-1}} > 0$, which implies that $J_0^{e,lower}$ is an increasing function of P_{N-1} . For this case, in order to minimize $J_0^{e,lower}$, P_{N-1} must be chosen as 0; i.e., $P_{N-1}^* = 0$. Then, for $\lambda > \frac{\sigma_{M_{N-1}}^2}{\sigma_{W_{N-1}}^2}$, we have the following:

$$\begin{aligned} \frac{\partial J_0^{e,lower}}{\partial P_{N-2}} &= \lambda \left(1 + \frac{\partial P_{N-1}}{\partial P_{N-2}}\right) + \sum_{i=N-2}^{N-1} \frac{\partial \Delta_i}{\partial P_{N-2}} \\ &= \lambda \left(1 + \frac{\partial P_{N-1}}{\partial P_{N-2}}\right) + \left(g^2 \left(1 + \frac{P_{N-1}}{\sigma_{W_{N-1}}^2}\right)^{-1} + 1\right) \frac{\partial \Delta_{N-2}}{\partial P_{N-2}} \\ &\quad - \left(\sigma_{V_{N-2}}^2 + g^2 \Delta_{N-2}\right) \left(1 + \frac{P_{N-1}}{\sigma_{W_{N-1}}^2}\right)^{-2} \frac{1}{\sigma_{W_{N-1}}^2} \frac{\partial P_{N-1}}{\partial P_{N-2}} \\ &\stackrel{(a)}{=} \lambda + \frac{\partial \Delta_{N-2}}{\partial P_{N-2}} (g^2 + 1) \\ &= \lambda - \left(\sigma_{V_{N-3}}^2 + g^2 \Delta_{N-3}\right) \left(1 + \frac{P_{N-2}}{\sigma_{W_{N-2}}^2}\right)^{-2} \frac{1}{\sigma_{W_{N-2}}^2} (g^2 + 1) \\ &\geq \lambda - \left(\sigma_{V_{N-3}}^2 + g^2 \sigma_{M_{N-3}}^2\right) \frac{1}{\sigma_{W_{N-2}}^2} (g^2 + 1) \\ &= \lambda - \frac{\sigma_{M_{N-2}}^2}{\sigma_{W_{N-2}}^2} (g^2 + 1). \end{aligned}$$

Here, (a) holds since $P_{N-1}^* = 0$ for $\lambda > \frac{\sigma_{M_{N-1}}^2}{\sigma_{W_{N-1}}^2}$. If $\lambda > \max \left\{ \frac{\sigma_{M_{N-1}}^2}{\sigma_{W_{N-1}}^2}, \frac{\sigma_{M_{N-2}}^2}{\sigma_{W_{N-2}}^2} (g^2 + 1) \right\}$, then $\frac{\partial J_0^{e,lower}}{\partial P_{N-2}} > 0$, which implies that $J_0^{e,lower}$ is an increasing function of P_{N-2} . For this case, in order to minimize

$J_0^{e,lower}$, P_{N-2} must be chosen as 0. By following the similar approach and assumptions on λ , since $P_{N-1}^* = P_{N-2}^* = \dots = P_{k+1}^* = 0$, we have the following:

$$\begin{aligned}
\frac{\partial J_0^{e,lower}}{\partial P_k} &= \lambda + \sum_{i=k}^{N-1} \frac{\partial \Delta_i}{\partial P_k} \\
&= \lambda + \frac{\partial \Delta_k}{\partial P_k} \sum_{i=k}^{N-1} \prod_{j=k+1}^i g^2 \\
&= \lambda - \left(\sigma_{V_{k-1}}^2 + g^2 \Delta_{k-1} \right) \left(1 + \frac{P_k}{\sigma_{W_k}^2} \right)^{-2} \frac{1}{\sigma_{W_k}^2} \sum_{i=k}^{N-1} \prod_{j=k+1}^i g^2 \\
&\geq \lambda - \left(\sigma_{V_{k-1}}^2 + g^2 \sigma_{M_{k-1}}^2 \right) \frac{1}{\sigma_{W_k}^2} \sum_{i=k}^{N-1} \prod_{j=k+1}^i g^2 \\
&= \lambda - \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \sum_{i=k}^{N-1} g^{2(i-k)} \\
&= \lambda - \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \sum_{i=0}^{N-k-1} g^{2i},
\end{aligned}$$

where $\prod_{i=k}^l = 1$ if $k > l$. If $\lambda > \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \sum_{i=0}^{N-k-1} g^{2i}$, then $\frac{\partial J_0^{e,lower}}{\partial P_k} > 0$, which implies that $J_0^{e,lower}$ is an increasing function of P_k . For this case, in order to minimize $J_0^{e,lower}$, P_k must be chosen as 0.

By combining all the results above, it can be deduced that if $\lambda > \max_{k=0,1,\dots,N-1} \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \sum_{i=0}^{N-k-1} g^{2i}$, then the lower bound $J_0^{e,lower}$ of the encoder costs J_0^e is minimized by choosing $P_0^* = P_1^* = \dots = P_{N-1}^* = 0$; that is, the encoder does not signal any output. Hence, the encoder engages in a non-informative equilibrium and the minimum cost becomes $J_0^e = J_0^{e,lower} = \left(\sum_{i=0}^{N-1} \sigma_{M_i}^2 \right) + Nb^2$ at this non-informative equilibrium. ■

Now consider the dynamic Stackelberg signaling game with a discounted infinite horizon and a discount factor $\beta \in (0, 1)$; i.e., $J^e(\gamma^e, \gamma^d) = \mathbb{E} \left[\sum_{i=0}^{\infty} \beta^i ((m_i - u_i - b)^2 + \lambda x_i^2) \right]$ and $J^d(\gamma^e, \gamma^d) = \mathbb{E} \left[\sum_{i=0}^{\infty} \beta^i (m_i - u_i)^2 \right]$.

Theorem 3.5: There does not exist an informative (affine or non-linear) equilibrium in the infinite horizon discounted dynamic Stackelberg signaling game for scalar Gauss-Markov sources; i.e., the only equilibrium is the non-informative one, if $\lambda \geq \max_{k=0,1,\dots} \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \frac{1}{1-\beta g^2}$ where $\beta g^2 < 1$.

Proof: For the infinite horizon case, it can be observed

$$\inf_{\gamma_{[0,N-1]}^e} \lim_{N \rightarrow \infty} J^e(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d) \geq \limsup_{N \rightarrow \infty} \inf_{\gamma_{[0,N-1]}^e} \sum_{i=0}^{N-1} \beta^i (\Delta_i + \lambda P_i + b^2).$$

Thus, $\limsup_{N \rightarrow \infty} \inf_{\gamma_{[0,N-1]}^e} \sum_{i=0}^{N-1} \beta^i (\Delta_i + \lambda P_i + b^2)$ is achieved at a non-informative equilibrium if $\lambda > \limsup_{N \rightarrow \infty} \max_{k=0,1,\dots,N-1} \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \sum_{i=0}^{N-k-1} \beta^i g^{2i} = \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \frac{1}{1-\beta g^2}$ for $\beta g^2 < 1$. Hence, if $\lambda \geq$

$\max_{k=0,1,\dots} \frac{\sigma_{M_k}^2}{\sigma_{W_k}^2} \frac{1}{1-\beta g^2}$, then the lower bound $J_0^{e,lower}$ of the encoder costs J_0^e is minimized by choosing $P_0 = P_1 = \dots = 0$, and the minimum cost becomes $J_0^e = J_0^{e,lower} = \sum_{i=0}^{\infty} \beta^i (\sigma_{M_i}^2 + b^2)$ at this non-informative equilibrium. ■

IV. DYNAMIC QUADRATIC SIGNALING GAMES FOR MULTI-DIMENSIONAL GAUSS-MARKOV SOURCES

In this section, the scalar setup considered in Section III is extended to the n -dimensional setup as follows: The source is assumed to be an n -dimensional Markovian source with initial Gaussian distribution; i.e., $\vec{M}_0 \sim \mathcal{N}(0, \Sigma_{\vec{M}_0})$ and $\vec{M}_{k+1} = G\vec{M}_k + \vec{V}_k$ where G is an $n \times n$ matrix and $\vec{V}_k \sim \mathcal{N}(0, \Sigma_{\vec{V}_k})$ is an i.i.d. Gaussian noise sequence for $k = 0, 1, \dots, N-2$. The channels between the encoder and the decoder are assumed to be i.i.d. additive Gaussian channels; i.e., $\vec{W}_k \sim \mathcal{N}(0, \Sigma_{\vec{W}_k})$, and \vec{W}_k and \vec{V}_l are independent for $k = 0, 1, \dots, N-1$ and $l = 0, 1, \dots, N-2$. In the k -th stage of the N -stage game, the encoder knows the values of $\mathcal{I}_k^e = \{\vec{m}_{[0,k]}, \vec{y}_{[0,k-1]}\}$ (a noiseless feedback channel is assumed) and the decoder knows the values of $\mathcal{I}_k^d = \{\vec{y}_{[0,k]}\}$ with $\vec{y}_k = \vec{x}_k + \vec{w}_k$. Thus, under the policies considered, $\vec{x}_k = \gamma_k^e(\mathcal{I}_k^e)$ and $\vec{u}_k = \gamma_k^d(\mathcal{I}_k^d)$. The encoder's goal is to minimize

$$J^e(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d) = \mathbb{E} \left[\sum_{k=0}^{N-1} c_k^e(\vec{m}_k, \vec{x}_k, \vec{u}_k) \right],$$

whereas, the decoder's goal is to minimize

$$J^d(\gamma_{[0,N-1]}^e, \gamma_{[0,N-1]}^d) = \mathbb{E} \left[\sum_{k=0}^{N-1} c_k^d(\vec{m}_k, \vec{u}_k) \right]$$

by finding the optimal policy sequences $\gamma_{[0,N-1]}^e$ and $\gamma_{[0,N-1]}^d$, respectively. The cost functions are $c^e(\vec{m}_k, \vec{x}_k, \vec{u}_k) = \|\vec{m}_k - \vec{u}_k - \vec{b}\|^2 + \lambda \|\vec{x}_k\|^2$ and $c^d(\vec{m}_k, \vec{u}_k) = \|\vec{m}_k - \vec{u}_k\|^2$ where the lengths of the vectors are defined in L_2 norm and \vec{b} is the bias vector.

A. Dynamic Nash Equilibria for Vector Gauss-Markov Sources

Similar to the scalar source case, affine policies constitute an invariant subspace under the best response maps for Nash equilibria when the source is multi-dimensional in the dynamic signaling games as shown below:

Theorem 4.1:

- i) If the encoder uses affine policies at all stages, then the decoder will be affine at all stages.
- ii) If the decoder uses affine policies at all stages, then the encoder will be affine at all stages.

Proof:

- i) Let the encoder policies be $\vec{x}_k = \gamma_k^e(\vec{m}_{[0,k]}, \vec{y}_{[0,k-1]}) = \sum_{i=0}^k A_{k,i} \vec{m}_i + \sum_{i=0}^{k-1} B_{k,i} \vec{y}_i + \vec{C}_k$ where $A_{k,i}$ and $B_{k,i}$ are $n \times n$ matrices, and C_k is $n \times 1$ vector for $k \leq N-1$ and $i \leq k$. Similar to the dynamic multi-dimensional Stackelberg cheap talk analysis in Theorem 2.8, the optimal decoder actions can be found as $\vec{u}_k^* = \gamma_k^{*,d}(\mathcal{I}_k^d) = \mathbb{E}[\vec{m}_k | \mathcal{I}_k^d] = \mathbb{E}[\vec{m}_k | \vec{y}_{[0,k]}]$ for $k \leq N-1$. Notice that $\vec{y}_{[0,k]}$ is multivariate Gaussian for $k \leq N-1$ since $\vec{y}_k = \vec{x}_k + \vec{w}_k$. This proves that $\gamma_k^{*,d}(\mathcal{I}_k^d)$ is an affine function of $\vec{y}_{[0,k]}$ due to the joint Gaussianity.
- ii) See Appendix C. ■

B. Dynamic Stackelberg Equilibria for Vector Gauss-Markov Sources

Even when the encoder and the decoder have identical (non-biased) quadratic cost functions, when the source and the channel are multi-dimensional, linear policies may not be optimal; see [4, Chapter 11] for a detailed discussion. In particular, except for settings where matching between the source and the channel exists (building on [29], [30]), the optimality of linear policies is quite rare [31]. Matching essentially requires that the capacity achieving source probabilities and the rate-distortion achieving channel probabilistic characteristics are simultaneously realized for a given system; this is precisely the case for a scalar Gaussian source transmitted over a scalar additive Gaussian channel. One special case where such a matching holds is the case when the noise and signal power levels are identical in every channel and the distortion criterion is identical for all scalar components [32]. For further discussions on multi-dimensional Gaussian source and channel pairs, we refer the reader to [30]–[37].

It is evident from Theorem 4.1 that when the encoder is linear, the optimal decoder is linear. In this case, a relevant problem is to find the optimal Stackelberg policy among the linear or affine class.

In the following, a dynamic programming approach is adapted to find such Stackelberg equilibria. Building on the optimality of linear innovation encoders, we restrict the analysis to such encoders. Our analysis builds on and generalizes the arguments in [38, Theorem 3] and [39].

Theorem 4.2: Suppose that G , $\Sigma_{\vec{M}_0}$ and $\Sigma_{\vec{v}_k}$ are diagonal. Suppose further that the innovation is given by $\tilde{\vec{m}}_k \triangleq \vec{m}_k - \mathbb{E}[\vec{m}_k | \vec{y}_{[0,k-1]}]$ with $\tilde{\vec{m}}_0 = \vec{m}_0$, and that the encoder linearly encodes the innovation. Then, an optimal such linear policy can be computed through dynamic programming with value functions $V_k(\Sigma_{\tilde{\vec{m}}_k}) \triangleq \text{tr}(K_k \Sigma_{\tilde{\vec{m}}_k} + L_k)$ that satisfy the terminal condition $V_N(\Sigma_{\tilde{\vec{m}}_N}) = 0$ with diagonal K_k matrices for $k = 0, 1, \dots, N-1$.

Proof: We will follow an approach similar to that in [38] which restricted the analysis to a team problem and a scalar channel; [38] in turn builds on [39], which considers continuous time systems. Since the $(k+1)$ st stage encoder policy only transmits the linearly encoded innovation by

assumption, $\vec{x}_k = \gamma_k^e(\mathcal{I}_k^e) = A_k \tilde{\vec{m}}_k$ where A_k is an $n \times n$ matrix for $k = 0, 1, \dots, N-1$. Then the decoder receives $\vec{y}_k = \vec{x}_k + \vec{w}_k = A_k \tilde{\vec{m}}_k + \vec{w}_k$ and applies the action $\vec{u}_k = \gamma_k^d(\mathcal{I}_k^d) = \mathbb{E}[\vec{m}_k | \vec{y}_{[0,k]}]$ to minimize his stage-wise cost $\|\vec{e}_k\|^2 \triangleq \mathbb{E}[\|\vec{m}_k - \vec{u}_k\|^2] = \mathbb{E}[(\vec{m}_k - \vec{u}_k)^T (\vec{m}_k - \vec{u}_k)] = \text{tr}(\Sigma_{\vec{e}_k})$ for $k = 0, 1, \dots, N-1$ where $\Sigma_{\vec{R}}$ stands for the covariance matrix of the random variable \vec{R} ; i.e., $\Sigma_{\vec{R}} \triangleq \mathbb{E}[(\vec{R} - \mathbb{E}[\vec{R}])(\vec{R} - \mathbb{E}[\vec{R}])^T]$. Due to the orthogonality of $\tilde{\vec{m}}_k$ and $\vec{y}_{[0,k-1]}$, and the iterated expectations rule, $\vec{u}_k = \mathbb{E}[\vec{m}_k | \vec{y}_{[0,k]}] = \mathbb{E}[\tilde{\vec{m}}_k + \mathbb{E}[\vec{m}_k | \vec{y}_{[0,k-1]}] | \vec{y}_{[0,k]}] = \mathbb{E}[\tilde{\vec{m}}_k | \vec{y}_k] + \mathbb{E}[\vec{m}_k | \vec{y}_{[0,k-1]}]$, and it follows that $\vec{e}_k = \vec{m}_k - \vec{u}_k = \vec{m}_k - \mathbb{E}[\tilde{\vec{m}}_k | \vec{y}_k] - \mathbb{E}[\vec{m}_k | \vec{y}_{[0,k-1]}] = \tilde{\vec{m}}_k - \mathbb{E}[\tilde{\vec{m}}_k | \vec{y}_k]$. Since $\mathbb{E}[\tilde{\vec{m}}_k | \vec{y}_k] = \Sigma_{\tilde{\vec{m}}_k} A_k^T (\Sigma_{\vec{y}_k})^{-1} \vec{y}_k$, the stage-wise cost of the decoder becomes the trace of the following:

$$\begin{aligned}
\Sigma_{\vec{e}_k} &= \mathbb{E}[\vec{e}_k \vec{e}_k^T] = \mathbb{E}[(\vec{m}_k - \vec{u}_k)(\vec{m}_k - \vec{u}_k)^T] \\
&= \mathbb{E}\left[\left(\tilde{\vec{m}}_k - \Sigma_{\tilde{\vec{m}}_k} A_k^T (\Sigma_{\vec{y}_k})^{-1} \vec{y}_k\right) \left(\tilde{\vec{m}}_k - \Sigma_{\tilde{\vec{m}}_k} A_k^T (\Sigma_{\vec{y}_k})^{-1} \vec{y}_k\right)^T\right] \\
&= \Sigma_{\tilde{\vec{m}}_k} - \Sigma_{\tilde{\vec{m}}_k} A_k^T (\Sigma_{\vec{y}_k})^{-1} A_k \Sigma_{\tilde{\vec{m}}_k} = \Sigma_{\tilde{\vec{m}}_k} - \Sigma_{\tilde{\vec{m}}_k} A_k^T \left(A_k \Sigma_{\tilde{\vec{m}}_k} A_k^T + \Sigma_{\vec{w}_k}\right)^{-1} A_k \Sigma_{\tilde{\vec{m}}_k} \\
&= \Sigma_{\tilde{\vec{m}}_k} - \Sigma_{\tilde{\vec{m}}_k} A_k^T \Sigma_{\vec{w}_k}^{-1/2} \left(\Sigma_{\vec{w}_k}^{-1/2} A_k \Sigma_{\tilde{\vec{m}}_k} A_k^T \Sigma_{\vec{w}_k}^{-1/2} + I\right)^{-1} \Sigma_{\vec{w}_k}^{-1/2} A_k \Sigma_{\tilde{\vec{m}}_k} \\
&= \Sigma_{\tilde{\vec{m}}_k}^{1/2} \left(I - \Sigma_{\tilde{\vec{m}}_k}^{1/2} A_k^T \Sigma_{\vec{w}_k}^{-1/2} \left(\Sigma_{\vec{w}_k}^{-1/2} A_k \Sigma_{\tilde{\vec{m}}_k}^{1/2} \Sigma_{\tilde{\vec{m}}_k}^{1/2} A_k^T \Sigma_{\vec{w}_k}^{-1/2} + I\right)^{-1} \Sigma_{\vec{w}_k}^{-1/2} A_k \Sigma_{\tilde{\vec{m}}_k}^{1/2}\right) \Sigma_{\tilde{\vec{m}}_k}^{1/2} \\
&\stackrel{(a)}{=} \Sigma_{\tilde{\vec{m}}_k}^{1/2} \left(I - H_k^T (H_k H_k^T + I)^{-1} H_k\right) \Sigma_{\tilde{\vec{m}}_k}^{1/2} \\
&\stackrel{(b)}{=} \Sigma_{\tilde{\vec{m}}_k}^{1/2} (I + H_k^T H_k)^{-1} \Sigma_{\tilde{\vec{m}}_k}^{1/2}, \tag{9}
\end{aligned}$$

where (a) follows from $H_k \triangleq \Sigma_{\vec{w}_k}^{-1/2} A_k \Sigma_{\tilde{\vec{m}}_k}^{1/2}$, and (b) follows from the matrix inversion lemma, $(I + UWV)^{-1} = I - U(W^{-1} + VU)^{-1}V$, by choosing $U = H_k^T$, $W = I$, and $V = H_k$.

Observe the following identity:

$$\begin{aligned}
\mathbb{E}[\vec{m}_k | \vec{y}_k] &= \mathbb{E}[\vec{m}_k] + \mathbb{E}[\vec{m}_k \vec{y}_k^T] (\Sigma_{\vec{y}_k})^{-1} \vec{y}_k = \mathbb{E}\left[(\tilde{\vec{m}}_k + \mathbb{E}[\vec{m}_k | \vec{y}_{[0,k-1]}])(A_k \tilde{\vec{m}}_k + \vec{w}_k)^T\right] (\Sigma_{\vec{y}_k})^{-1} \vec{y}_k \\
&= \left(\Sigma_{\tilde{\vec{m}}_k} A_k^T + \mathbb{E}\left[\mathbb{E}[\vec{m}_k | \vec{y}_{[0,k-1]}] \tilde{\vec{m}}_k^T\right] A_k^T\right) (\Sigma_{\vec{y}_k})^{-1} \vec{y}_k \\
&\stackrel{(a)}{=} \Sigma_{\tilde{\vec{m}}_k} A_k^T (\Sigma_{\vec{y}_k})^{-1} \vec{y}_k,
\end{aligned}$$

where (a) due to the orthogonality of $\mathbb{E}[\vec{m}_k | \vec{y}_{[0,k-1]}]$ and $\tilde{\vec{m}}_k$. Then the innovation can be expressed recursively as follows:

$$\begin{aligned}
\tilde{\vec{m}}_{k+1} &= \vec{m}_{k+1} - \mathbb{E}[\vec{m}_{k+1} | \vec{y}_{[0,k]}] \\
&= G\vec{m}_k + \vec{v}_k - \mathbb{E}[\vec{m}_{k+1} | \vec{y}_{[0,k-1]}] - \mathbb{E}[\vec{m}_{k+1} | \vec{y}_k] \\
&= G\vec{m}_k + \vec{v}_k - G\mathbb{E}[\vec{m}_k | \vec{y}_{[0,k-1]}] - G\mathbb{E}[\vec{m}_k | \vec{y}_k] \\
&= G\tilde{\vec{m}}_k + \vec{v}_k - G\mathbb{E}[\vec{m}_k | \vec{y}_k]
\end{aligned}$$

$$= G\tilde{\vec{m}}_k + \vec{v}_k - G\Sigma_{\tilde{\vec{m}}_k} A_k^T (\Sigma_{\vec{y}_k})^{-1} \vec{y}_k.$$

Then the covariance matrices of the innovations can be expressed as

$$\begin{aligned} \Sigma_{\tilde{\vec{m}}_{k+1}} &= \mathbb{E}[\tilde{\vec{m}}_{k+1} \tilde{\vec{m}}_{k+1}^T] \\ &= G\Sigma_{\tilde{\vec{m}}_k} G^T + \Sigma_{\vec{v}_k} - G\Sigma_{\tilde{\vec{m}}_k} A_k^T (\Sigma_{\vec{y}_k})^{-1} A_k \Sigma_{\tilde{\vec{m}}_k} G^T \\ &= G\Sigma_{\tilde{\vec{m}}_k} G^T + \Sigma_{\vec{v}_k} - G\Sigma_{\tilde{\vec{m}}_k} A_k^T \left(A_k \Sigma_{\tilde{\vec{m}}_k} A_k^T + \Sigma_{\vec{w}_k} \right)^{-1} A_k \Sigma_{\tilde{\vec{m}}_k} G^T \\ &= G\Sigma_{\tilde{\vec{m}}_k}^{1/2} (I + H_k^T H_k)^{-1} \Sigma_{\tilde{\vec{m}}_k}^{1/2} G^T + \Sigma_{\vec{v}_k}. \end{aligned} \quad (10)$$

The optimal encoder chooses A_k in order to minimize his stage-wise cost

$$\begin{aligned} \mathbb{E} \left[\|\vec{m}_k - \vec{u}_k - \vec{b}\|^2 + \lambda \|\vec{x}_k\|^2 \right] &= \mathbb{E} \left[\|\vec{m}_k - \mathbb{E}[\vec{m}_k | \vec{y}_{[0,k]}] - \vec{b}\|^2 + \lambda \|\vec{x}_k\|^2 \right] \\ &= \mathbb{E} \left[\|\vec{m}_k - \mathbb{E}[\vec{m}_k | \vec{y}_{[0,k]}]\|^2 + \|\vec{b}\|^2 + \lambda \|\vec{x}_k\|^2 \right] \\ &= \text{tr}(\Sigma_{\vec{e}_k}) + \text{tr}(\lambda \Sigma_{\vec{x}_k}) + \|\vec{b}\|^2 \\ &\stackrel{(a)}{=} \text{tr} \left(\Sigma_{\tilde{\vec{m}}_k}^{1/2} (I + H_k^T H_k)^{-1} \Sigma_{\tilde{\vec{m}}_k}^{1/2} \right) + \text{tr} \left(\lambda A_k \Sigma_{\tilde{\vec{m}}_k} A_k^T \right) + \|\vec{b}\|^2 \\ &= \text{tr} \left(\Sigma_{\tilde{\vec{m}}_k} (I + H_k^T H_k)^{-1} \right) + \text{tr} (\lambda H_k^T \Sigma_{\vec{w}_k} H_k) + \|\vec{b}\|^2 \end{aligned} \quad (11)$$

where (a) is obtained by using (9).

Let the value functions be $V_k(\Sigma_{\tilde{\vec{m}}_k}) = \text{tr}(K_k \Sigma_{\tilde{\vec{m}}_k} + L_k)$ with K_k being diagonal. In the following we show that there exist such V_k that satisfy Bellman's principle of optimality [40, Theorem 3.2.1]. Here $V_k(\Sigma_{\tilde{\vec{m}}_k}) \triangleq \min_{H_k} \left(\mathcal{C}_k(\Sigma_{\tilde{\vec{m}}_k}, H_k) + V_{k+1}(\Sigma_{\tilde{\vec{m}}_{k+1}}) \right)$ and $\mathcal{C}_k(\Sigma_{\tilde{\vec{m}}_k}, H_k) \triangleq \text{tr}(\Sigma_{\vec{e}_k}) + \text{tr}(\lambda \Sigma_{\vec{x}_k}) + \|\vec{b}\|^2$ is the stage-wise cost of the k -th stage encoder. Then,

$$\begin{aligned} V_k(\Sigma_{\tilde{\vec{m}}_k}) &= \min_{H_k} \left(\mathcal{C}_k(\Sigma_{\tilde{\vec{m}}_k}, H_k) + V_{k+1}(\Sigma_{\tilde{\vec{m}}_{k+1}}) \right) \\ &\stackrel{(a)}{=} \min_{H_k} \left(\text{tr} \left(\Sigma_{\tilde{\vec{m}}_k} (I + H_k^T H_k)^{-1} \right) + \text{tr} (\lambda H_k^T \Sigma_{\vec{w}_k} H_k) + \|\vec{b}\|^2 + \text{tr} (K_{k+1} \Sigma_{\tilde{\vec{m}}_{k+1}} + L_{k+1}) \right) \\ &\stackrel{(b)}{=} \min_{H_k} \left(\text{tr} \left(\Sigma_{\tilde{\vec{m}}_k} (I + H_k^T H_k)^{-1} \right) + \text{tr} (\lambda H_k^T \Sigma_{\vec{w}_k} H_k) + \|\vec{b}\|^2 \right. \\ &\quad \left. + \text{tr} \left(K_{k+1} G \Sigma_{\tilde{\vec{m}}_k}^{1/2} (I + H_k^T H_k)^{-1} \Sigma_{\tilde{\vec{m}}_k}^{1/2} G^T + K_{k+1} \Sigma_{\vec{v}_k} + L_{k+1} \right) \right) \\ &\stackrel{(c)}{=} \text{tr} (K_{k+1} \Sigma_{\vec{v}_k} + L_{k+1}) + \|\vec{b}\|^2 \\ &\quad + \min_{H_k} \left(\text{tr} \left(\Sigma_{\tilde{\vec{m}}_k}^{1/2} (I + H_k^T H_k)^{-1} \Sigma_{\tilde{\vec{m}}_k}^{1/2} \right) + \text{tr} (\lambda H_k^T \Sigma_{\vec{w}_k} H_k) \right) \end{aligned}$$

$$\begin{aligned}
& + \text{tr} \left(G^T K_{k+1} G \Sigma_{\tilde{m}_k}^{1/2} (I + H_k^T H_k)^{-1} \Sigma_{\tilde{m}_k}^{1/2} \right) \\
& = \text{tr} (K_{k+1} \Sigma_{\tilde{v}_k} + L_{k+1}) + \|\vec{b}\|^2 \\
& + \min_{H_k} \left(\text{tr} \left(\Sigma_{\tilde{m}_k}^{1/2} (G^T K_{k+1} G + I) \Sigma_{\tilde{m}_k}^{1/2} (I + H_k^T H_k)^{-1} \right) + \text{tr} (\lambda H_k^T \Sigma_{\tilde{w}_k} H_k) \right) \quad (12)
\end{aligned}$$

where (a) follows by substituting $\mathcal{C}_k(\Sigma_{\tilde{m}_k}, H_k)$ using (11), (b) follows by employing (10), and (c) follows from the fact that K_{k+1} and L_{k+1} do not depend on H_k . The equivalent problem of the minimization of $\text{tr} \left(\Sigma_{\tilde{m}_k}^{1/2} (G^T K_{k+1} G + I) \Sigma_{\tilde{m}_k}^{1/2} (I + H_k^T H_k)^{-1} \right)$ over H_k under the constraint $\text{tr} (\lambda H_k^T \Sigma_{\tilde{w}_k} H_k) = \mu_k$ is considered in [41], and the solution technique can be adapted as follows:

Let $\nu_{k_1} \geq \nu_{k_2} \geq \dots \geq \nu_{k_n} > 0$ and $\tau_{k_n} \geq \tau_{k_{n-1}} \geq \dots \geq \tau_{k_1} > 0$ be the eigenvalues of $\Sigma_{\tilde{m}_k}^{1/2} (G^T K_{k+1} G + I) \Sigma_{\tilde{m}_k}^{1/2}$ and $\lambda \Sigma_{\tilde{w}_k}$, respectively, and $\mu_{k_p} \triangleq \sum_{i=1}^p (\sqrt{\tau_{k_i} \nu_{k_i}} - \tau_{k_i})$. If μ_{k_p} values are non-positive for $p = 1, 2, \dots, n$, then the optimal H_k becomes zero; i.e., $H_k^* = 0$. Otherwise, check if $\tau_{k_p}/\nu_{k_p} < 1$ for p that makes μ_{k_p} positive. If the inequality is not satisfied, again the optimal H_k becomes zero; i.e., $H_k^* = 0$. Finally, pick p and corresponding μ_{k_p} which give the minimum of $\frac{(\sum_{i=1}^p \sqrt{\tau_{k_i} \nu_{k_i}})^2}{\mu_{k_p} + \sum_{i=1}^p \tau_{k_i}} + \sum_{i=p+1}^n \nu_{k_i} + \mu_{k_p}$. If that minimum is greater than $\text{tr} \left(\Sigma_{\tilde{m}_k}^{1/2} (G^T K_{k+1} G + I) \Sigma_{\tilde{m}_k}^{1/2} \right)$, then the optimal H_k becomes zero; i.e., $H_k^* = 0$. Otherwise, the optimal H_k is found as $H_k^* = \Pi_k \zeta_k P_k^T$ where Π_k is a unitary matrix such that $\Pi_k^T (\lambda \Sigma_{\tilde{w}_k}) \Pi_k = \text{diag}(\tau_{k_1}, \tau_{k_2}, \dots, \tau_{k_n}) \triangleq \tilde{\Pi}_k$, P_k is a unitary matrix such that $P_k^T \left(\Sigma_{\tilde{m}_k}^{1/2} (G^T K_{k+1} G + I) \Sigma_{\tilde{m}_k}^{1/2} \right) P_k = \text{diag}(\nu_{k_1}, \nu_{k_2}, \dots, \nu_{k_n})$, and ζ_k is a diagonal matrix such that $\zeta_k = \text{diag}(\sqrt{\alpha_{k_1}}, \sqrt{\alpha_{k_2}}, \dots, \sqrt{\alpha_{k_{p_k^*}}}, 0, \dots, 0)$ with $\alpha_{k_i} = -1 + \left(\frac{\sqrt{\nu_{k_i}/\tau_{k_i}}}{\sum_{j=1}^{p_k^*} \sqrt{\tau_{k_j} \nu_{k_j}}} \right) \left(1 + \sum_{j=1}^{p_k^*} \tau_{k_j} \right)$. For $p \leq n$, let $f_k(p) \triangleq \sqrt{\tau_{k_p}/\nu_{k_p}} \sum_{i=1}^p \sqrt{\tau_{k_i} \nu_{k_i}} - \sum_{i=1}^p \tau_{k_i}$, then p_k^* is defined by

$$p_k^* = n \text{ if } f_k(n) < 1$$

$$f_k(p_k^*) < 1 \leq f_k(p_k^* + 1) \text{ if } f_k(n) \geq 1.$$

Since the optimal H_k always has the form of $H_k^* = \Pi_k \zeta_k P_k^T$ for every $\mu_k = \text{tr} (\lambda H_k^T \Sigma_{\tilde{w}_k} H_k)$ as described above, then the recursion of the innovation's covariance matrix (10) can be expressed as

$$\Sigma_{\tilde{m}_{k+1}} = G \Sigma_{\tilde{m}_k}^{1/2} (I + P_k \zeta_k^T \zeta_k P_k^T)^{-1} \Sigma_{\tilde{m}_k}^{1/2} G^T + \Sigma_{\tilde{v}_k} \quad (13)$$

Then (12) becomes

$$\begin{aligned}
V_k(\Sigma_{\tilde{m}_k}) & = \text{tr} (K_{k+1} \Sigma_{\tilde{v}_k} + L_{k+1}) + \|\vec{b}\|^2 \\
& + \text{tr} \left(\Sigma_{\tilde{m}_k}^{1/2} (G^T K_{k+1} G + I) \Sigma_{\tilde{m}_k}^{1/2} (I + P_k \zeta_k^T \Pi_k^T \Pi_k \zeta_k P_k^T)^{-1} \right) + \text{tr} (\lambda P_k \zeta_k^T \Pi_k^T \Sigma_{\tilde{w}_k} \Pi_k \zeta_k P_k^T)
\end{aligned}$$

$$\begin{aligned}
&\stackrel{(a)}{=} \text{tr} (K_{k+1} \Sigma_{\vec{v}_k} + L_{k+1}) + \|\vec{b}\|^2 \\
&\quad + \text{tr} \left(\Sigma_{\tilde{m}_k}^{1/2} (G^T K_{k+1} G + I) \Sigma_{\tilde{m}_k}^{1/2} (I + P_k \zeta_k^T \zeta_k P_k^T)^{-1} \right) + \text{tr} \left(\zeta_k^T \tilde{\Pi}_k \zeta_k \right) \\
&\stackrel{(b)}{=} \text{tr} (K_{k+1} \Sigma_{\vec{v}_k} + L_{k+1}) + \|\vec{b}\|^2 \\
&\quad + \text{tr} \left(\Sigma_{\tilde{m}_k}^{1/2} (G^T K_{k+1} G + I) \Sigma_{\tilde{m}_k}^{1/2} \left(I - P_k \zeta_k^T (I + \zeta_k P_k^T P_k \zeta_k^T)^{-1} \zeta_k P_k^T \right) \right) + \text{tr} \left(\zeta_k^T \tilde{\Pi}_k \zeta_k \right) \\
&\stackrel{(c)}{=} \text{tr} (K_{k+1} \Sigma_{\vec{v}_k} + L_{k+1}) + \|\vec{b}\|^2 \\
&\quad + \text{tr} \left(\Sigma_{\tilde{m}_k}^{1/2} (G^T K_{k+1} G + I) \Sigma_{\tilde{m}_k}^{1/2} \left(I - P_k \zeta_k^T (I + \zeta_k \zeta_k^T)^{-1} \zeta_k P_k^T \right) \right) + \text{tr} \left(\zeta_k^T \tilde{\Pi}_k \zeta_k \right) \\
&\stackrel{(d)}{=} \text{tr} (K_{k+1} \Sigma_{\vec{v}_k} + L_{k+1}) + \|\vec{b}\|^2 + \text{tr} \left(\zeta_k^T \tilde{\Pi}_k \zeta_k \right) \\
&\quad + \text{tr} \left((G^T K_{k+1} G + I) \left(\Sigma_{\tilde{m}_k} - \zeta_k^T (I + \zeta_k \zeta_k^T)^{-1} \zeta_k \Sigma_{\tilde{m}_k} \right) \right) \\
&= \text{tr} (K_{k+1} \Sigma_{\vec{v}_k} + L_{k+1}) + \text{tr} (\vec{b} \vec{b}^T) + \text{tr} \left(\zeta_k^T \tilde{\Pi}_k \zeta_k \right) \\
&\quad + \text{tr} \left((G^T K_{k+1} G + I) \left(I - \zeta_k^T (I + \zeta_k \zeta_k^T)^{-1} \zeta_k \right) \Sigma_{\tilde{m}_k} \right) \tag{14}
\end{aligned}$$

where (a) follows from $\Pi_k^T \Pi_k = I$, $\tilde{\Pi}_k = \Pi_k^T (\lambda \Sigma_{\vec{w}_k}) \Pi_k$, and the properties of the trace operator, (b) follows from the matrix inversion lemma $(I + UWV)^{-1} = I - U(W^{-1} + VU)^{-1}V$, by choosing $U = P_k \zeta_k^T$, $W = I$, and $V = \zeta_k P_k^T$, (c) is due to $P_k^T P_k = I$, and (d) follows from the diagonality of $\Sigma_{\tilde{m}_k}$, P_k and ζ_k : Since G , $\Sigma_{\tilde{m}_0}$, K_k and $\Sigma_{\vec{v}_k}$ are diagonal for $k = 0, 1, \dots, N-1$, it is always possible to find a unitary diagonal P_0 such that $P_0^T \left(\Sigma_{\tilde{m}_0}^{1/2} (G^T K_1 G + I) \Sigma_{\tilde{m}_0}^{1/2} \right) P_0 = \text{diag}(\nu_{01}, \nu_{02}, \dots, \nu_{0n})$, which makes $\Sigma_{\tilde{m}_1}$ diagonal by (13). By following the same approach, $\Sigma_{\tilde{m}_k}$ and P_k are diagonal for $k = 0, 1, \dots, N-1$.

In order to satisfy (14), since $V_N \left(\Sigma_{\tilde{m}_N} \right) = 0$, we choose $K_N = L_N = 0$, and $\{K_{k+1}, L_{k+1}\}$ according to

$$\begin{aligned}
K_k &= \left(G^T K_{k+1} G + I \right) \left(I - \zeta_k^T (I + \zeta_k \zeta_k^T)^{-1} \zeta_k \right) \\
L_k &= K_{k+1} \Sigma_{\vec{v}_k} + L_{k+1} + \zeta_k^T \tilde{\Pi}_k \zeta_k + \vec{b} \vec{b}^T. \tag{15}
\end{aligned}$$

for $k = 0, 1, \dots, N-1$. Now we verify that the diagonal K_k matrices satisfy the dynamic programming recursion. ■

For the special case when the channel is scalar, this result reduces to

$$\begin{aligned}
K_k &= \left(G^T K_{k+1} G + I \right) \times \text{diag} \left(\frac{\lambda \sigma_{W_k}^2}{1 + \lambda \sigma_{W_k}^2}, 1, 1, \dots, 1 \right) \\
L_k &= K_{k+1} \Sigma_{\vec{v}_k} + L_{k+1} + \text{diag} (1, 0, 0, \dots, 0) + \vec{b} \vec{b}^T
\end{aligned}$$

for $k = 0, 1, \dots, N - 1$. The optimal linear encoder policy is found as $A_k^* = \Sigma_{\tilde{w}_k}^{1/2} \zeta_k P_k^T \Sigma_{\tilde{m}_k}^{-1/2}$ since the optimal H_k^* is $H_k^* = \Pi_k \zeta_k P_k^T$, the H_k is defined as $H_k \triangleq \Sigma_{\tilde{w}_k}^{-1/2} A_k \Sigma_{\tilde{m}_k}^{1/2}$, and $\Pi_k = 1$ and $\zeta_k = \left[\frac{1}{\sqrt{\lambda \sigma_{W_k}^2}}, 0, \dots, 0 \right]$ for the scalar channel.

V. CONCLUDING REMARKS

In this paper, Nash and Stackelberg equilibria for dynamic quadratic cheap talk and signaling games have been analyzed. For the cheap talk problem under Nash equilibria, we have shown that the last stage equilibria are quantized for any scalar source with an arbitrary distribution, and fully revealing equilibria cannot exist in general (see Remark 2.1) whereas for the dynamic Stackelberg cheap talk game, the equilibria must be fully revealing regardless of the source model. We have also proved that the equilibria are fully revealing in the dynamic multi-dimensional cheap talk under Stackelberg equilibria whereas the equilibria cannot be fully revealing under a Nash concept. In the dynamic signaling game where the transmission of a Gaussian source over a Gaussian channel is considered, affine policies constitute an invariant subspace under best response maps for scalar and multi-dimensional sources under Nash equilibria. However, for dynamic Stackelberg signaling games involving Gauss-Markov sources and memoryless Gaussian channels, we have proved that for scalar setups linear policies are optimal and the only equilibrium is the linear one, whereas this is not the case for general multi-dimensional setups. Finally, the conditions under which the equilibrium is non-informative under the Stackelberg assumption are derived for scalar Gauss-Markov sources, and the dynamic programming formulation is presented for a class of Stackelberg equilibria when the encoders are restricted to be linear for multi-dimensional Gauss-Markov sources.

APPENDIX A

PROOF OF THEOREM 3.2

Recall that in the static Nash signaling game, the optimal affine encoder cost is obtained as $J^{*,e} = \sigma_M^2 + b^2$ for $\lambda \geq \sigma_M^2/\sigma_W^2$ (in non-informative equilibrium) and $J^{*,e} = 2\sqrt{\lambda\sigma_M^2\sigma_W^2} + b^2\sqrt{\frac{\sigma_M^2}{\lambda\sigma_W^2}} - \lambda\sigma_W^2$ for $\lambda < \sigma_M^2/\sigma_W^2$ (in informative affine equilibrium) [2]. Even though $\sigma_M^2 + b^2 \geq 2\sqrt{\lambda\sigma_M^2\sigma_W^2} + b^2\sqrt{\frac{\sigma_M^2}{\lambda\sigma_W^2}} - \lambda\sigma_W^2$ for $\lambda \geq \sigma_M^2/\sigma_W^2$ (since $\left(\sqrt{\sigma_M^2} - \sqrt{\lambda\sigma_W^2}\right)^2 \geq 0$), the non-informative equilibrium is preferred over the informative affine equilibrium since there does not exist any informative affine equilibrium for $\lambda \geq \sigma_M^2/\sigma_W^2$. However, the non-informative equilibrium can also be preferred even if $\lambda < \sigma_M^2/\sigma_W^2$.

Next, we analyze the conditions under which the informative affine equilibrium has strictly lower cost than the non-informative one for $\lambda < \sigma_M^2/\sigma_W^2$:

$$\begin{aligned} \sigma_M^2 + b^2 &> 2\sqrt{\lambda\sigma_M^2\sigma_W^2} + b^2\sqrt{\frac{\sigma_M^2}{\lambda\sigma_W^2}} - \lambda\sigma_W^2 \\ \Rightarrow \left(\sqrt{\sigma_M^2} - \sqrt{\lambda\sigma_W^2}\right) \left(\sqrt{\lambda\sigma_W^2} \left(\sqrt{\sigma_M^2} - \sqrt{\lambda\sigma_W^2}\right) - b^2\right) &> 0 \end{aligned} \quad (16)$$

Since $\lambda < \sigma_M^2/\sigma_W^2$, (16) reduces to $\sqrt{\lambda\sigma_W^2} \left(\sqrt{\sigma_M^2} - \sqrt{\lambda\sigma_W^2}\right) - b^2 > 0$. Let $t \triangleq \sqrt{\lambda\sigma_W^2}$, then the inequality becomes $t^2 - \sqrt{\sigma_M^2}t + b^2 < 0$. If $\sigma_M^2 < 4b^2$, then $t^2 - \sqrt{\sigma_M^2}t + b^2$ is always positive, which implies that the non-informative equilibrium has strictly less cost than the informative affine equilibrium. Otherwise; i.e., if $\sigma_M^2 \geq 4b^2$, then $t^2 - \sqrt{\sigma_M^2}t + b^2 < 0$ holds when $\frac{\sqrt{\sigma_M^2} - \sqrt{\sigma_M^2 - 4b^2}}{2} < t < \frac{\sqrt{\sigma_M^2} + \sqrt{\sigma_M^2 - 4b^2}}{2}$. Then, by inserting $t \triangleq \sqrt{\lambda\sigma_W^2}$,

$$0 < \frac{\sigma_M^2 - 2b^2 - \sqrt{\sigma_M^2}\sqrt{\sigma_M^2 - 4b^2}}{2\sigma_W^2} < \lambda < \frac{\sigma_M^2 - 2b^2 + \sqrt{\sigma_M^2}\sqrt{\sigma_M^2 - 4b^2}}{2\sigma_W^2} < \frac{\sigma_M^2}{\sigma_W^2}.$$

Thus, the encoder prefers the informative affine equilibrium when $\frac{\sigma_M^2 - 2b^2 - \sqrt{\sigma_M^2}\sqrt{\sigma_M^2 - 4b^2}}{2\sigma_W^2} < \lambda < \frac{\sigma_M^2 - 2b^2 + \sqrt{\sigma_M^2}\sqrt{\sigma_M^2 - 4b^2}}{2\sigma_W^2}$ and $\sigma_M^2 \geq 4b^2$; otherwise, the non-informative equilibrium is preferred.

APPENDIX B

PROOF OF THEOREM 3.3

If the decoder policies are $u_0 = \gamma_0^d(y_0) = Ky_0 + L$ and $u_1 = \gamma_1^d(y_0, y_1) = M_0y_0 + M_1y_1 + N$ where K, L, M_0, M_1 and N are scalars, then the optimal encoder policies are as follows:

$$\begin{aligned} \gamma_0^{*,e}(m_0) &= \frac{KM_1^2 + \lambda K + \lambda g M_0}{(\lambda + K^2)(\lambda + M_1^2) + \lambda M_0^2} m_0 - \frac{K(M_1^2 + \lambda)(L + b) + \lambda M_0(N + b)}{(\lambda + K^2)(\lambda + M_1^2) + \lambda M_0^2} \\ \gamma_1^{*,e}(m_0, m_1, y_0) &= \frac{M_1}{M_1^2 + \lambda} m_1 - \frac{M_0 M_1}{M_1^2 + \lambda} y_0 - \frac{M_1(N + b)}{M_1^2 + \lambda} \end{aligned}$$

If the encoder policies are $x_0 = \gamma_0^e(m_0) = Am_0 + C$ and $x_1 = \gamma_1^e(m_0, m_1, y_0) = Fm_0 + B_0y_0 + B_1m_1 + D$ where A, C, F, B_0, B_1 and D are scalars, then the optimal decoder policies are as follows:

$$\begin{aligned} \gamma_0^{*,d}(y_0) &= \frac{A\sigma_{M_0}^2}{A^2\sigma_{M_0}^2 + \sigma_{W_0}^2} y_0 - \frac{AC\sigma_{M_0}^2}{A^2\sigma_{M_0}^2 + \sigma_{W_0}^2} \\ \gamma_1^{*,d}(y_0, y_1) &= \frac{-(g^2B_0B_1 + gB_0F)\sigma_{M_0}^2\sigma_{W_0}^2 - (A^2B_0B_1 + AB_1F)\sigma_{M_0}^2\sigma_{V_0}^2 + gA\sigma_{M_0}^2\sigma_{W_1}^2 - B_0B_1\sigma_{V_0}^2\sigma_{W_0}^2}{(gB_1 + F)^2\sigma_{M_0}^2\sigma_{W_0}^2 + A^2\sigma_{M_0}^2\sigma_{W_1}^2 + A^2B_1^2\sigma_{M_0}^2\sigma_{V_0}^2 + \sigma_{W_0}^2\sigma_{W_1}^2 + B_1^2\sigma_{V_0}^2\sigma_{W_0}^2} y_0 \\ &\quad + \frac{(g^2B_1 + gF)\sigma_{M_0}^2\sigma_{W_0}^2 + A^2B_1\sigma_{M_0}^2\sigma_{V_0}^2 + B_1\sigma_{V_0}^2\sigma_{W_0}^2}{(gB_1 + F)^2\sigma_{M_0}^2\sigma_{W_0}^2 + A^2\sigma_{M_0}^2\sigma_{W_1}^2 + A^2B_1^2\sigma_{M_0}^2\sigma_{V_0}^2 + \sigma_{W_0}^2\sigma_{W_1}^2 + B_1^2\sigma_{V_0}^2\sigma_{W_0}^2} y_1 \\ &\quad - \frac{gAC\sigma_{M_0}^2\sigma_{W_1}^2 + (g^2B_1 + gF)D\sigma_{M_0}^2\sigma_{W_0}^2 + (A^2B_1D - AB_1CF)\sigma_{M_0}^2\sigma_{V_0}^2 + B_1D\sigma_{V_0}^2\sigma_{W_0}^2}{(gB_1 + F)^2\sigma_{M_0}^2\sigma_{W_0}^2 + A^2\sigma_{M_0}^2\sigma_{W_1}^2 + A^2B_1^2\sigma_{M_0}^2\sigma_{V_0}^2 + \sigma_{W_0}^2\sigma_{W_1}^2 + B_1^2\sigma_{V_0}^2\sigma_{W_0}^2} \end{aligned}$$

After some algebraic manipulations, consider the following cases:

- 1) $\frac{(g^2 + 1)\sigma_{M_0}^2/\sigma_{W_0}^2 < \sigma_{M_1}^2/\sigma_{W_1}^2}{A}$: By using the relation between A and K , and assuming nonzero A ,

$$M_1^2 = \frac{\lambda a^2 K^2 \sigma_{W_0}^2}{\lambda A^2 \sigma_{M_0}^2 - K^2 \sigma_{W_0}^2} - \lambda \geq 0 \Rightarrow \lambda \leq (g^2 + 1) \frac{\sigma_{M_0}^2}{\sigma_{W_0}^2}$$

is obtained. Hence, for $\lambda > (g^2 + 1) \frac{\sigma_{M_0}^2}{\sigma_{W_0}^2}$, there does not exist an informative affine equilibrium in the first stage since $A = K = 0$.

Now suppose that $\lambda > (g^2 + 1) \frac{\sigma_{M_0}^2}{\sigma_{W_0}^2}$. Then the two-stage game setup reduces to the one-stage game setup; i.e., $A = C = B_0 = K = L = M_0 = 0$, and the encoder and the decoder policies are $\gamma_1^e(m_1) = B_1 m_1 + D$ and $\gamma_1^d(y_1) = M_1 y_1 + N$, respectively. Thus, the equilibrium is informative if and only if $\max \left\{ \frac{\sigma_{M_1}^2 - 2b^2 - \sqrt{\sigma_{M_1}^2 \sigma_{W_1}^2 - 4b^2}}{2\sigma_{W_1}^2}, (g^2 + 1) \frac{\sigma_{M_0}^2}{\sigma_{W_0}^2} \right\} < \lambda < \frac{\sigma_{M_1}^2 - 2b^2 + \sqrt{\sigma_{M_1}^2 \sigma_{W_1}^2 - 4b^2}}{2\sigma_{W_1}^2}$ and $\sigma_{M_1}^2 \geq 4b^2$ by Theorem 3.2.

- 2) $\frac{\sigma_{M_1}^2/\sigma_{W_1}^2 < (g^2 + 1)\sigma_{M_0}^2/\sigma_{W_0}^2}{M_1}$: By using the relation between M_1 and B_1 , and assuming nonzero M_1 ,

$$M_1^2 = \sqrt{\frac{\lambda}{\sigma_{W_1}^2} \left(g^2 \frac{\sigma_{M_0}^2 \sigma_{W_0}^2}{A^2 \sigma_{M_0}^2 + \sigma_{W_0}^2} + \sigma_{V_0}^2 \right)} - \lambda \geq 0 \Rightarrow \lambda \leq \frac{g^2 \sigma_{M_0}^2 + \sigma_{V_0}^2}{\sigma_{W_1}^2} = \frac{\sigma_{M_1}^2}{\sigma_{W_1}^2}$$

is obtained. Hence, for $\lambda > \frac{\sigma_{M_1}^2}{\sigma_{W_1}^2}$, $M_1 = 0$ and the second stage message m_1 will not be used in the game.

Now suppose that $\lambda > \frac{\sigma_{M_1}^2}{\sigma_{W_1}^2}$. Then $B_0 = B_1 = D = M_1 = 0$, $M_0 = gK$, $N = gL$, $C = -\frac{b}{\lambda}(g+1)K$, $L = \frac{b}{\lambda}(g+1)K^2$, and $A = \frac{(g^2+1)K}{(g^2+1)K^2+\lambda}$. By using the relation between A and K , and assuming nonzero K , $((g^2 + 1)K^2 + \lambda)^2 \sigma_{W_0}^2 = \lambda(g^2 + 1)\sigma_{M_0}^2$ is obtained, which implies $K = \sqrt{\sqrt{\frac{\lambda}{g^2+1} \frac{\sigma_{M_0}^2}{\sigma_{W_0}^2}} - \frac{\lambda}{g^2+1}}$ and $A = \sqrt{\sqrt{\frac{g^2+1}{\lambda} \frac{\sigma_{W_0}^2}{\sigma_{M_0}^2}} - \frac{\sigma_{W_0}^2}{\sigma_{M_0}^2}}$ for $0 < \lambda \leq (g^2 + 1) \frac{\sigma_{M_0}^2}{\sigma_{W_0}^2}$. Hence, if $\frac{\sigma_{M_1}^2}{\sigma_{W_1}^2} < \lambda \leq \frac{(g^2+1)\sigma_{M_0}^2}{\sigma_{W_0}^2}$, the second stage message m_1 will not be used in the game. If $\lambda > \max \left\{ \frac{(g^2+1)\sigma_{M_0}^2}{\sigma_{W_0}^2}, \frac{\sigma_{M_1}^2}{\sigma_{W_1}^2} \right\}$, then $A = K = L = C = N = M_0 = 0$, which leads to a non-informative equilibrium.

APPENDIX C

PROOF OF THEOREM 4.1

Let the decoder policies be $\vec{u}_k = \gamma_k^d(\vec{y}_{[0,k]}) = \sum_{i=0}^k K_{k,i} \vec{y}_i + \vec{L}_k$ where $K_{k,i}$ is $n \times n$ matrix and \vec{L}_k is $n \times 1$ vector for $k \leq N-1$ and $i \leq k$. With $\vec{y}_{N-1} = \vec{x}_{N-1} + \vec{w}_{N-1}$, it follows that

$\vec{u}_{N-1} = \sum_{i=0}^{N-2} K_{N-1,i} \vec{y}_i + K_{N-1,N-1} \vec{x}_{N-1} + K_{N-1,N-1} \vec{w}_{N-1} + \vec{L}_{N-1}$. Then, by a dynamic programming approach, the final stage encoder cost can be written as

$$\begin{aligned}
J_{N-1}^{*,e} &= \min_{\vec{x}_{N-1} = \gamma_{N-1}^e(\vec{m}_{[0,N-1]}, \vec{y}_{[0,N-2]})} \mathbb{E} \left[\|\vec{m}_{N-1} - \vec{u}_{N-1} - \vec{b}\|^2 + \lambda \|\vec{x}_{N-1}\|^2 \right] \\
&\stackrel{(a)}{=} \min_{\vec{x}_{N-1}} \mathbb{E} \left[\left((K_{N-1,N-1}^T K_{N-1,N-1} + \lambda I) \vec{x}_{N-1} - K_{N-1,N-1}^T \left(\vec{m}_{N-1} - \sum_{i=0}^{N-2} K_{N-1,i} \vec{y}_i - \vec{L}_{N-1} - \vec{b} \right) \right)^T \right. \\
&\quad \times \left(K_{N-1,N-1}^T K_{N-1,N-1} + \lambda I \right)^{-1} \\
&\quad \times \left((K_{N-1,N-1}^T K_{N-1,N-1} + \lambda I) \vec{x}_{N-1} - K_{N-1,N-1}^T \left(\vec{m}_{N-1} - \sum_{i=0}^{N-2} K_{N-1,i} \vec{y}_i - \vec{L}_{N-1} - \vec{b} \right) \right) \\
&\quad \left. + \left(\vec{m}_{N-1} - \sum_{i=0}^{N-2} K_{N-1,i} \vec{y}_i - \vec{L}_{N-1} - \vec{b} \right)^T \left(I - K_{N-1,N-1} (K_{N-1,N-1}^T K_{N-1,N-1} + \lambda I)^{-1} K_{N-1,N-1}^T \right) \right. \\
&\quad \left. \times \left(\vec{m}_{N-1} - \sum_{i=0}^{N-2} K_{N-1,i} \vec{y}_i - \vec{L}_{N-1} - \vec{b} \right) + \vec{w}_{N-1}^T K_{N-1,N-1}^T K_{N-1,N-1} \vec{w}_{N-1} \right]
\end{aligned}$$

where (a) is obtained by completing the square. Hence, the optimal $\gamma_{N-1}^e(\vec{m}_{[0,N-1]}, \vec{y}_{[0,N-2]})$ is

$$\gamma_{N-1}^e(\vec{m}_{[0,N-1]}, \vec{y}_{[0,N-2]}) = \left(K_{N-1,N-1}^T K_{N-1,N-1} + \lambda I \right)^{-1} K_{N-1,N-1}^T \left(\vec{m}_{N-1} - \sum_{i=0}^{N-2} K_{N-1,i} \vec{y}_i - \vec{L}_{N-1} - \vec{b} \right)$$

and the minimum final stage encoder cost is obtained as

$$\begin{aligned}
J_{N-1}^{*,e} &= \mathbb{E} \left[\left(\vec{m}_{N-1} - \sum_{i=0}^{N-2} K_{N-1,i} \vec{y}_i - \vec{L}_{N-1} - \vec{b} \right)^T \left(I - K_{N-1,N-1} (K_{N-1,N-1}^T K_{N-1,N-1} + \lambda I)^{-1} K_{N-1,N-1}^T \right) \right. \\
&\quad \times \left. \left(\vec{m}_{N-1} - \sum_{i=0}^{N-2} K_{N-1,i} \vec{y}_i - \vec{L}_{N-1} - \vec{b} \right) + \vec{w}_{N-1}^T K_{N-1,N-1}^T K_{N-1,N-1} \vec{w}_{N-1} \right] \quad (17)
\end{aligned}$$

Notice that even though $\vec{m}_{[0,N-1]}$ and $\vec{y}_{[0,N-2]}$ are available at the encoder, the encoder uses only \vec{m}_{N-1} and $\vec{y}_{[0,N-2]}$ at the final stage; i.e., the encoder does not need $\vec{m}_{[0,N-2]}$.

Then, by a dynamic programming approach, the cost of the encoder at $N-1$ st stage can be written as

$$J_{N-2}^{*,e} = \min_{\vec{x}_{N-2} = \gamma_{N-2}^e(\vec{m}_{[0,N-2]}, \vec{y}_{[0,N-3]})} \mathbb{E} \left[\|\vec{m}_{N-2} - \vec{u}_{N-2} - \vec{b}\|^2 + \lambda \|\vec{x}_{N-2}\|^2 + J_{N-1}^{*,e} \right] \quad (18)$$

By using the relation between the sources $\vec{m}_{N-1} = G\vec{m}_{N-2} + \vec{v}_{N-2}$ and $\vec{y}_{N-2} = \vec{x}_{N-2} + \vec{w}_{N-2}$, and defining $\Omega_{N-1} = (I - K_{N-1,N-1} (K_{N-1,N-1}^T K_{N-1,N-1} + \lambda I)^{-1} K_{N-1,N-1}^T)$, (17) can be refined and inserted into (18). Further, with $\vec{y}_{N-2} = \vec{x}_{N-2} + \vec{w}_{N-2}$, it follows that $\vec{u}_{N-2} = \sum_{i=0}^{N-3} K_{N-2,i} \vec{y}_i +$

$K_{N-2,N-2} \vec{x}_{N-2} + K_{N-2,N-2} \vec{w}_{N-2} + \vec{L}_{N-2}$ and the completion of the squares method can be applied similar to the previous step. Then, the optimal $\gamma_{N-2}^e(\vec{m}_{[0,N-2]}, \vec{y}_{[0,N-3]})$ is obtained as

$$\begin{aligned} \gamma_{N-2}^e(\vec{m}_{[0,N-2]}, \vec{y}_{[0,N-3]}) = & \left(K_{N-2,N-2}^T K_{N-2,N-2} + \lambda I + K_{N-1,N-2}^T \Omega_{N-1} K_{N-1,N-2} \right)^{-1} \\ & \times \left(K_{N-2,N-2}^T \left(\vec{m}_{N-2} - \sum_{i=0}^{N-3} K_{N-2,i} \vec{y}_i - \vec{L}_{N-2} - \vec{b} \right) \right. \\ & \left. + K_{N-1,N-2}^T \Omega_{N-1} \left(G \vec{m}_{N-2} - \sum_{i=0}^{N-3} K_{N-1,i} \vec{y}_i - \vec{L}_{N-1} - \vec{b} \right) \right). \end{aligned}$$

Notice that even though $\vec{m}_{[0,N-2]}$ and $\vec{y}_{[0,N-3]}$ are available at the encoder, the encoder uses only \vec{m}_{N-2} and $\vec{y}_{[0,N-3]}$ at the final stage; i.e., the encoder does not need $\vec{m}_{[0,N-3]}$.

It is observed that the optimal \vec{x}_k can be obtained as an affine function of \vec{m}_k and $\vec{y}_{[0,k-1]}$ for each stage, $k = 0, 1, \dots, N-1$ by completing the square, since the cost of the current stage and the next stages consist of the quadratic function of \vec{x}_k after using the proper identities; i.e., $\vec{m}_k = G \vec{m}_{k-1} + \vec{v}_{k-1}$ and $\vec{y}_k = \vec{x}_k + \vec{w}_k$.

REFERENCES

- [1] V. P. Crawford and J. Sobel, "Strategic information transmission," *Econometrica*, vol. 50, pp. 1431–1451, 1982.
- [2] S. Saritaş, S. Yüksel, and S. Gezici, "Quadratic multi-dimensional signaling games and affine equilibria," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 605–619, Feb. 2017.
- [3] D. Blackwell, "The comparison of experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 93–102, 1951.
- [4] S. Yüksel and T. Başar, *Stochastic Networked Control Systems: Stabilization and Optimization under Information Constraints*. Boston, MA: Birkhäuser, 2013.
- [5] J. Hirshleifer, "The private and social value of information and the reward to inventive activity," *The American Economic Review*, pp. 561–574, 1971.
- [6] O. Gossner and J.-F. Mertens, "The value of information in zero-sum games," *preprint*, 2001.
- [7] M. I. Kamien, Y. Tauman, and S. Zamir, "On the value of information in a strategic conflict," *Games and Economic Behavior*, vol. 2, no. 2, pp. 129–153, 1990.
- [8] T. Başar, "Stochastic differential games and intricacy of information structures," in *Dynamic Games in Economics*, ser. Dynamic Modeling and Econometrics in Economics and Finance, J. Haunschild, V. M. Veliov, and S. Wrzaczek, Eds. Springer Berlin Heidelberg, 2014, vol. 16, pp. 23–49.
- [9] I. Shames, A. Teixeira, H. Sandberg, and K. Johansson, "Agents misbehaving in a network: a vice or a virtue?" *IEEE Network*, vol. 26, no. 3, pp. 35–40, May 2012.
- [10] J. Miklós-Thal and H. Schumacher, "The value of recommendations," *Games and Economic Behavior*, vol. 79, pp. 132–147, 2013.
- [11] B. Larrousse, O. Beaude, and S. Lasaulce, "Crawford-Sobel meet Lloyd-Max on the grid," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 6127–6131.
- [12] M. Golosov, V. Skreta, A. Tsyvinski, and A. Wilson, "Dynamic strategic information transmission," *Journal of Economic Theory*, vol. 151, pp. 304–341, 2014.

- [13] F. Farokhi, A. M. H. Teixeira, and C. Langbort, “Estimation with strategic sensors,” *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 724–739, Feb. 2017.
- [14] E. Akyol, C. Langbort, and T. Başar, “Information-theoretic approach to strategic communication as a hierarchical game,” *Proceedings of the IEEE*, vol. 105, no. 2, pp. 205–218, Feb. 2017.
- [15] M. O. Sayın, E. Akyol, and T. Başar, “Hierarchical multi-stage Gaussian signaling games,” *arXiv preprint arXiv:1609.09448*, 2016.
- [16] M. Le Treust and T. Tomala, “Information Design for Strategic Coordination of Autonomous Devices with Non-Aligned Utilities,” in *54th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, Illinois, United States, Sep. 2016.
- [17] V. Kavitha, E. Altman, R. El-Azouzi, and R. Sundaresan, “Opportunistic scheduling in cellular systems in the presence of noncooperative mobiles,” *IEEE Transactions on Information Theory*, vol. 58, no. 3, pp. 1757–1773, March 2012.
- [18] D. Vasal and A. Anastasopoulos, “Signaling equilibria for dynamic lqg games with asymmetric information,” in *2016 IEEE 55th Conference on Decision and Control (CDC)*, Dec 2016, pp. 6901–6908.
- [19] T. Başar and G. Olsder, *Dynamic Noncooperative Game Theory*. Philadelphia, PA: SIAM Classics in Applied Mathematics, 1999.
- [20] Y. Shoham and K. Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. New York, NY, USA: Cambridge University Press, 2009.
- [21] D. W. Yeung and L. A. Petrosjan, *Cooperative Stochastic Differential Games*. New York, NY, USA: Springer-Verlag New York, 2006.
- [22] S. Fabricius, P. Furrer, S. Kerner, T. Linder, and S. Yüksel, “Game theory and information, Queen’s University, MTHE 493 Technical Report,” Apr. 2014.
- [23] I. I. Gihman and A. V. Skorohod, *Controlled Stochastic Processes*. New York, NY: Springer-Verlag New York, 1979.
- [24] V. S. Borkar, “White-noise representations in stochastic realization theory,” *SIAM J. on Control and Optimization*, vol. 31, pp. 1093–1102, 1993.
- [25] D. Blackwell, “Memoryless strategies in finite-stage dynamic programming,” *Annals of Mathematical Statistics*, vol. 35, pp. 863–865, 1964.
- [26] D. Blackwell and C. Ryll-Nardzewski, “Non-existence of everywhere proper conditional distributions,” *Annals of Mathematical Statistics*, pp. 223–225, 1963.
- [27] R. Bansal and T. Başar, “Simultaneous design of measurement and control strategies for stochastic systems with feedback,” *Automatica*, vol. 45, no. 5, pp. 679–694, 1989.
- [28] —, “Simultaneous design of measurement and control strategies in stochastic systems with feedback,” *Automatica*, vol. 45, pp. 679–694, September 1989.
- [29] C. D. Charalambous, P. A. Stavrou, and N. U. Ahmed, “Nonanticipative rate distortion function and relations to filtering theory,” *IEEE Transactions on Automatic Control*, vol. 59, no. 4, pp. 937–952, Apr. 2014.
- [30] M. Gastpar, B. Rimoldi, and M. Vetterli, “To code, or not to code: Lossy source-channel communication revisited,” *IEEE Transactions on Information Theory*, vol. 49, pp. 1147–1158, May 2003.
- [31] E. Akyol and K. Rose, “On linear transforms in zero-delay Gaussian source channel coding,” in *Proceedings of the IEEE International Symposium on Information Theory*, Boston, MA, 2012, pp. 1548–1552.
- [32] R. Pilc, “The optimum linear modulator for a Gaussian source used with a Gaussian channel,” *IEEE Transactions on Automatic Control*, vol. 48, pp. 3075–3089, Nov. 1969.
- [33] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Channels*. Budapest: Akademiai Kiado, 1981.
- [34] T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1971.
- [35] S. Tatikonda and S. Mitter, “Control under communication constraints,” *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1056–1068, 2004.

- [36] S. Tatikonda, A. Sahai, and S. Mitter, “Stochastic linear control over a communication channels,” *IEEE Transactions on Automatic Control*, vol. 49, pp. 1549–1561, Sep. 2004.
- [37] K. H. Lee and D. P. Petersen, “Optimal linear coding for vector channels,” *IEEE Transactions on Communications*, vol. 24, pp. 1283–1290, 1976.
- [38] A. A. Zaidi, T. J. Oechtering, S. Yüksel, and M. Skoglund, “Stabilization and control over Gaussian networks,” in *Information and Control in Networks*, Editors: G. Como, B. Bernhardsson, A. Rantzer. Springer, 2013.
- [39] T. Başar and R. Bansal, “Optimum design of measurement channels and control policies for linear-quadratic stochastic systems,” *European J. Operations Research*, vol. 73, pp. 226–236, Dec. 1994.
- [40] O. Hernandez-Lerma and J. Lasserre, *Discrete-time Markov control processes*. Springer, 1996.
- [41] T. Başar, “A trace minimization problem with applications in joint estimation and control under nonclassical information,” *J. of Optimization Theory and Applications*, vol. 31, pp. 343–359, July 1980.